

La Courciere  
08/765244 Page 1  
Seq. IDs 1-22 w/  
Interf

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 18, 2004, 02:41:57 ; Search time 324.095 Seconds  
(without alignments)  
140.169 Million cell updates/sec

Title: US-08-765-244-1  
Perfect score: 208  
Sequence: 1 MLSNLRILLNKALRKAKHTS.....VNFYRGKPVQSLKPRDLIC 41

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 6730630 seqs, 1107998698 residues

Total number of hits satisfying chosen parameters: 6730630

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending Patents AA Main:\*

- 1: /cgn2\_6/ptodata/1/paa/PCTUS COMB.pcp.\*
- 2: /cgn2\_6/ptodata/1/paa/US06 COMB.pcp.\*
- 3: /cgn2\_6/ptodata/1/paa/US07 COMB.pcp.\*
- 4: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 5: /cgn2\_6/ptodata/1/paa/US081 COMB.pcp.\*
- 6: /cgn2\_6/ptodata/1/paa/US082 COMB.pcp.\*
- 7: /cgn2\_6/ptodata/1/paa/US083 COMB.pcp.\*
- 8: /cgn2\_6/ptodata/1/paa/US084 COMB.pcp.\*
- 9: /cgn2\_6/ptodata/1/paa/US085 COMB.pcp.\*
- 10: /cgn2\_6/ptodata/1/paa/US086 COMB.pcp.\*
- 11: /cgn2\_6/ptodata/1/paa/US087 COMB.pcp.\*
- 12: /cgn2\_6/ptodata/1/paa/US088 COMB.pcp.\*
- 13: /cgn2\_6/ptodata/1/paa/US089 COMB.pcp.\*
- 14: /cgn2\_6/ptodata/1/paa/US090 COMB.pcp.\*
- 15: /cgn2\_6/ptodata/1/paa/US091 COMB.pcp.\*
- 16: /cgn2\_6/ptodata/1/paa/US092 COMB.pcp.\*
- 17: /cgn2\_6/ptodata/1/paa/US093 COMB.pcp.\*
- 18: /cgn2\_6/ptodata/1/paa/US094 COMB.pcp.\*
- 19: /cgn2\_6/ptodata/1/paa/US095 COMB.pcp.\*
- 20: /cgn2\_6/ptodata/1/paa/US096 COMB.pcp.\*
- 21: /cgn2\_6/ptodata/1/paa/US097A COMB.pcp.\*
- 22: /cgn2\_6/ptodata/1/paa/US097B COMB.pcp.\*
- 23: /cgn2\_6/ptodata/1/paa/US098 COMB.pcp.\*
- 24: /cgn2\_6/ptodata/1/paa/US099A COMB.pcp.\*
- 25: /cgn2\_6/ptodata/1/paa/US099B COMB.pcp.\*
- 26: /cgn2\_6/ptodata/1/paa/US100 COMB.pcp.\*
- 27: /cgn2\_6/ptodata/1/paa/US101 COMB.pcp.\*
- 28: /cgn2\_6/ptodata/1/paa/US102 COMB.pcp.\*
- 29: /cgn2\_6/ptodata/1/paa/US103 COMB.pcp.\*
- 30: /cgn2\_6/ptodata/1/paa/US104 COMB.pcp.\*
- 31: /cgn2\_6/ptodata/1/paa/US105 COMB.pcp.\*
- 32: /cgn2\_6/ptodata/1/paa/US106 COMB.pcp.\*
- 33: /cgn2\_6/ptodata/1/paa/US107 COMB.pcp.\*
- 34: /cgn2\_6/ptodata/1/paa/US108 COMB.pcp.\*
- 35: /cgn2\_6/ptodata/1/paa/US109 COMB.pcp.\*
- 36: /cgn2\_6/ptodata/1/paa/US160 COMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

\*

Result No.	Score	Match	Length	DB	ID	Description
1	208	100.0	41	11	US-08-765-244-1	Sequence 1, Appl
2	197	94.7	43	11	US-08-765-244-22	Sequence 22, Appl
3	188	90.4	354	22	US-09-791-537-57942	Sequence 57942, A
4	179	86.1	354	22	US-09-791-537-17653	Sequence 17653, A
5	179	86.1	354	22	US-09-791-537-84892	Sequence 84892, A
6	179	86.1	354	28	US-10-219-051B-12767	Sequence 12767, A
7	179	86.1	354	28	US-10-219-051B-14340	Sequence 14340, A
8	144	69.2	354	22	US-09-791-537-27693	Sequence 27693, A
9	144	69.2	354	22	US-09-791-537-104643	Sequence 104643, A
10	144	69.2	366	1	PCT-US04-20454-90	Sequence 90, Appl
11	125	60.1	354	1	PCT-US04-20454-89	Sequence 89, Appl
12	125	60.1	354	21	US-09-724-676-76239	Sequence 76239, A
13	125	60.1	354	21	US-09-724-676A-76239	Sequence 76239, A
14	125	60.1	354	22	US-09-791-537-2694	Sequence 2694, Ap
15	125	60.1	354	22	US-09-791-537-46324	Sequence 46324, A
16	125	60.1	354	22	US-09-791-537-88687	Sequence 88687, A
17	125	60.1	354	28	US-10-219-051B-12769	Sequence 12769, A
18	125	60.1	354	28	US-10-219-051B-14342	Sequence 14342, A
19	114	54.8	354	22	US-09-791-537-53980	Sequence 53980, A
20	92	44.2	32	1	PCT-US02-02814-48	Sequence 48, Appl
21	92	44.2	32	1	PCT-US02-21677-54	Sequence 54, Appl
22	92	44.2	32	1	PCT-US04-11481-23	Sequence 23, Appl
23	92	44.2	32	1	PCT-US04-11482-23	Sequence 23, Appl
24	92	44.2	32	1	PCT-US04-23191-26	Sequence 26, Appl
25	92	44.2	32	1	PCT-US04-23192-26	Sequence 26, Appl
26	92	44.2	32	3	US-07-916-939-29	Sequence 29, Appl
27	92	44.2	32	8	US-08-438-190-29	Sequence 29, Appl
28	92	44.2	32	16	US-09-287-145-29	Sequence 29, Appl
29	92	44.2	32	19	US-09-522-727B-29	Sequence 29, Appl
30	92	44.2	32	19	US-09-522-727D-29	Sequence 29, Appl
31	92	44.2	32	26	US-10-052-942-54	Sequence 54, Appl
32	92	44.2	32	26	US-10-061-395-48	Sequence 48, Appl
33	92	44.2	32	34	US-10-823-254-23	Sequence 23, Appl
34	92	44.2	32	34	US-10-823-259-23	Sequence 23, Appl
35	92	44.2	354	22	US-09-791-537-53995	Sequence 53995, A
36	60	28.8	205	16	US-09-270-849B-181269	Sequence 181269, A
37	58	27.9	99	23	US-09-855-604-639	Sequence 639, App
38	58	27.9	99	23	US-09-855-604A-639	Sequence 85254, A
39	57	27.4	350	22	US-09-791-537-85254	Sequence 85254, A
40	56	26.9	302	29	US-10-369-493-3761	Sequence 3761, Ap
41	56	26.9	302	36	US-60-360-039-3761	Sequence 3761, Ap
42	56	26.9	356	22	US-09-791-537-53985	Sequence 53985, A
43	56	26.9	1235	21	US-09-708-427-25599	Sequence 25599, A
44	56	26.9	1247	21	US-09-708-427-25598	Sequence 25598, A
45	56	26.9	1286	1	PCT-US02-18153-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1  
US-08-765-244-1  
; Sequence 1, Application US/08765244  
; GENERAL INFORMATION:  
; APPLICANT: Seibel, Peter  
; APPLICANT: Seibel, Andrea  
; TITLE OF INVENTION: CHIMICAL PEPTIDE-NUCLEIC ACID  
; TITLE OF INVENTION: FRAGMENT, PROCESS FOR PRODUCING THE SAME AND ITS USES FOR  
; TITLE OF INVENTION: APPROPRIATELY INTRODUCING NUCLEIC ACIDS INTO CELL ORGANELLES  
; TITLE OF INVENTION: AND CELLS  
; FILE REFERENCE: 8484-0018-999  
; CURRENT APPLICATION NUMBER: US/08/765,244  
; CURRENT FILING DATE: 1997-10-30  
; PRIOR APPLICATION NUMBER: PCT/DE95/00775  
; PRIOR FILING DATE: 1995-06-11  
; PRIOR APPLICATION NUMBER: DE P 44 21 079.5  
; PRIOR FILING DATE: 1994-06-16  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1

;  
; LENGTH: 41  
; TYPE: PRT  
; ORGANISM: Rattus rattus  
; FEATURE:  
US-08-765-244-1

Query Match 100.0%; Score 208; DB 11; Length 41;  
Best Local Similarity 100.0%; Pred. No. 7.8e-24;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLSNLRILLNKAALRKAHTSMVNFYRGKPVQSQOLKPRDL 41  
DB 1 MLSNLRILLNKAALRKAHTSMVNFYRGKPVQSQOLKPRDL 41

## RESULT 2

US-08-765-244-22  
; Sequence 22, Application US/08765244  
; GENERAL INFORMATION:  
; APPLICANT: Seibel, Peter  
; APPLICANT: Seibel, Andrea  
; TITLE OF INVENTION: CHIMERICAL PEPTIDE-NUCLEIC ACID  
; TITLE OF INVENTION: FRAGMENT PROCESS FOR PRODUCING THE SAME AND ITS USES FOR  
; TITLE OF INVENTION: APPROPRIATELY INTRODUCING NUCLEIC ACIDS INTO CELL ORGANELLES  
; TITLE OF INVENTION: AND CELLS  
; FILE REFERENCE: 8484-0018-999  
; CURRENT APPLICATION NUMBER: US/08/765,244  
; CURRENT FILING DATE: 1997-10-30  
; PRIOR APPLICATION NUMBER: PCT/DE95/00775  
; PRIOR FILING DATE: 1995-06-11  
; PRIOR APPLICATION NUMBER: DE P 44 21 079.5  
; PRIOR FILING DATE: 1994-06-16  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 22  
; LENGTH: 43  
; TYPE: PRT  
; ORGANISM: Rattus rattus  
US-08-765-244-22

Query Match 94.7%; Score 197; DB 11; Length 43;  
Best Local Similarity 95.3%; Pred. No. 4.2e-22;  
Matches 41; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

QY 1 MLSNLRILLNKAALRKAHTSMVNFYRGKPVQSQOLKPRDL 41  
DB 1 MLSNLRILLNKAALRKAHTSMVNFYRGKPVQSQOLKPRDL 43

## RESULT 3

US-09-791-537-57942  
; Sequence 57942, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
; TITLE OF INVENTION: METHODS OF USE THEREOF  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 57942  
; LENGTH: 354  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-09-791-537-57942

Query Match 90.4%; Score 188; DB 22; Length 354;  
Best Local Similarity 95.2%; Pred. No. 1.5e-19;  
Matches 40; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

QY 1 MLSNLRILLNKAALRKAHTSMVNFYRGKPVQSQOLKPRDL 40  
DB 1 MLSNLRILLNKAALRKAHTSMVNFYRGKPVQSQOLKPRDL 42

## RESULT 4

US-09-791-537-17653  
; Sequence 17653, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
; TITLE OF INVENTION: METHODS OF USE THEREOF  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 17653  
; LENGTH: 354  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-09-791-537-17653

Query Match 86.1%; Score 179; DB 22; Length 354;  
Best Local Similarity 92.9%; Pred. No. 3.8e-18;  
Matches 39; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

QY 1 MLSNLRILLNKAALRKAHTSMVNFYRGKPVQSQOLKPRDL 40  
DB 1 MLSNLRILLNKAALRKAHTSMVNFYRGKPVQSQOLKPRDL 42

## RESULT 5

US-09-791-537-84892  
; Sequence 84892, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
; TITLE OF INVENTION: METHODS OF USE THEREOF  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 84892  
; LENGTH: 354  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-09-791-537-84892

Query Match 86.1%; Score 179; DB 22; Length 354;  
Best Local Similarity 92.9%; Pred. No. 3.8e-18;  
Matches 39; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

QY 1 MLSNLRILLNKAALRKAHTSMVNFYRGKPVQSQOLKPRDL 40  
DB 1 MLSNLRILLNKAALRKAHTSMVNFYRGKPVQSQOLKPRDL 42

## RESULT 6

US-10-219-051B-12767  
; Sequence 12767, Application US/10219051B  
; GENERAL INFORMATION:  
; APPLICANT: The General Hospital Corporation doing business as Massachusetts General  
; APPLICANT: Hospital / Bayer AG  
; TITLE OF INVENTION: Nucleotide sequences involved in pain  
; FILE REFERENCE: Lea 35693 Foreign Countries  
; CURRENT APPLICATION NUMBER: US/10/219,051B  
; CURRENT FILING DATE: 2003-05-09

; PRIOR APPLICATION NUMBER: US 60/312,147  
; PRIOR FILING DATE: 2001-08-14  
; PRIOR APPLICATION NUMBER: US 60/346,382  
; PRIOR FILING DATE: 2001-11-01  
; PRIOR APPLICATION NUMBER: US 60/333,347  
; PRIOR FILING DATE: 2001-11-26  
; NUMBER OF SEQ ID NOS: 14715  
; SOFTWARE: Perl script  
; SEQ ID NO 12767  
; LENGTH: 354  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
; PUBLICATION INFORMATION:  
; DATABASE ACCESSION NUMBER: SWISS-Prot / AAA41767  
; DATABASE ENTRY DATE: 2001-10-16  
US-10-219-051B-12767

Query Match 86.1%; Score 179; DB 28; Length 354;

Best Local Similarity 92.9%; Pred. No. 3.8e-18;  
Matches 39; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

QY 1 MLSNRLILNKAALRKAHTSMVRNFRYKPVQS--OLKPRDL 40  
|||||  
Db 1 MLSNRLILNKAALRKAHTSMVRNFRYKPVQSQVLKGRDL 42

## RESULT 7

US-10-219-051B-14340  
; Sequence 14340, Application US/10219051B  
; GENERAL INFORMATION:  
; APPLICANT: The General Hospital Corporation doing business as Massachusetts General  
; APPLICANT: Hospital / Bayer AG  
; TITLE OF INVENTION: Nucleotide sequences involved in pain  
; FILE REFERENCE: Lea 35693 Foreign Countries  
; CURRENT APPLICATION NUMBER: US/10/219,051B  
; CURRENT FILING DATE: 2003-05-09  
; PRIOR APPLICATION NUMBER: US 60/312,147  
; PRIOR FILING DATE: 2001-08-14  
; PRIOR APPLICATION NUMBER: US 60/346,382  
; PRIOR FILING DATE: 2001-11-01  
; PRIOR APPLICATION NUMBER: US 60/333,347  
; PRIOR FILING DATE: 2001-11-26  
; NUMBER OF SEQ ID NOS: 14715  
; SOFTWARE: Perl script  
; SEQ ID NO 14340  
; LENGTH: 354  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
; PUBLICATION INFORMATION:  
; DATABASE ACCESSION NUMBER: SWISS-Prot / OMWT  
; DATABASE ENTRY DATE: 2001-10-16  
US-10-219-051B-14340

Query Match 86.1%; Score 179; DB 28; Length 354;

Best Local Similarity 92.9%; Pred. No. 3.8e-18;  
Matches 39; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

QY 1 MLSNRLILNKAALRKAHTSMVRNFRYKPVQS--OLKPRDL 40  
|||||  
Db 1 MLSNRLILNKAALRKAHTSMVRNFRYKPVQSQVLKGRDL 42

## RESULT 8

US-09-791-537-27693  
; Sequence 27693, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; TITLE OF INVENTION: Danzer, Joseph  
; FILE REFERENCE: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537

; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 27693  
; LENGTH: 354  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-791-537-27693

Query Match 69.2%; Score 144; DB 22; Length 354;

Best Local Similarity 78.6%; Pred. No. 9.7e-13;  
Matches 33; Conservative 2; Mismatches 5; Indels 2; Gaps 1;

QY 1 MLSNRLILNKAALRKAHTSMVRNFRYKPVQS--OLKPRDL 40  
|||||  
Db 1 MLSNRLILNKAALRKAHTSMVRNFRYKPVQSQVLKGRDL 42

## RESULT 9

US-09-791-537-104643  
; Sequence 104643, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 104643  
; LENGTH: 354  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-791-537-104643

Query Match 69.2%; Score 144; DB 22; Length 354;

Best Local Similarity 78.6%; Pred. No. 9.7e-13;  
Matches 33; Conservative 2; Mismatches 5; Indels 2; Gaps 1;

QY 1 MLSNRLILNKAALRKAHTSMVRNFRYKPVQS--OLKPRDL 40  
|||||  
Db 1 MLSNRLILNKAALRKAHTSMVRNFRYKPVQSQVLKGRDL 42

## RESULT 10

PCT-US04-20454-90  
; Sequence 90, Application PC/TUS0420454  
; GENERAL INFORMATION:  
; APPLICANT: Gencia Corporation  
; APPLICANT: Khan, Shaharyar  
; TITLE OF INVENTION: Modified Vectors for Organelle Transfection  
; FILE REFERENCE: 120701-2020  
; CURRENT APPLICATION NUMBER: PCT/US04/20454  
; PRIOR FILING DATE: 2004-07-02  
; PRIOR APPLICATION NUMBER: 60/482,603  
; PRIOR FILING DATE: 2003-06-25  
; NUMBER OF SEQ ID NOS: 585  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 90  
; LENGTH: 366  
; TYPE: PRT  
; ORGANISM: Mus musculus  
PCT-US04-20454-90

Query Match 69.2%; Score 144; DB 1; Length 366;

Best Local Similarity 78.6%; Pred. No. 1e-12;  
Matches 33; Conservative 2; Mismatches 5; Indels 2; Gaps 1;

QY 1 MLSNRLILNKAALRKAHTSMVRNFRYKPVQS--OLKPRDL 40  
|||||





```
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 88687
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-88687

Query Match          60.1%; Score 125; DB 22; Length 354;
Best Local Similarity 66.7%; Pred. No. 8.4e-10;
Matches 28; Conservative 3; Mismatches 9; Indels 2; Gaps 1;

QY 1 MLSNRLILNKAALRKAHTSMVRNFRYKGPVQS--QLKPRDL 40
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MLFNLRLILNNAAFRNGHFMVVRNFRCGQPLQNKVKLGKRD 42
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 17
US-10-219-051B-12769
; Sequence 12769, Application US/10219051B
; GENERAL INFORMATION:
; APPLICANT: The General Hospital Corporation doing business as Massachusetts General
; APPLICANT: Hospital / Bayer AG
; TITLE OF INVENTION: Nucleotide sequences involved in pain
; FILE REFERENCE: Lea 35693 Foreign Countries
; CURRENT APPLICATION NUMBER: US/10/219,051B
; CURRENT FILING DATE: 2003-05-09
; PRIOR APPLICATION NUMBER: US 60/312,147
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/346,382
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/333,347
; PRIOR FILING DATE: 2001-11-26
; NUMBER OF SEQ ID NOS: 14715
; SOFTWARE: Perl script
; SEQ ID NO 12769
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: Refseq / NP_000522
; DATABASE ENTRY DATE: 2000-10-31
US-10-219-051B-12769

Query Match          60.1%; Score 125; DB 28; Length 354;
Best Local Similarity 66.7%; Pred. No. 8.4e-10;
Matches 28; Conservative 3; Mismatches 9; Indels 2; Gaps 1;

QY 1 MLSNRLILNKAALRKAHTSMVRNFRYKGPVQS--QLKPRDL 40
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MLFNLRLILNNAAFRNGHFMVVRNFRCGQPLQNKVKLGKRD 42
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 18
US-10-219-051B-14342
; Sequence 14342, Application US/10219051B
; GENERAL INFORMATION:
; APPLICANT: The General Hospital Corporation doing business as Massachusetts General
; APPLICANT: Hospital / Bayer AG
; TITLE OF INVENTION: Nucleotide sequences involved in pain
; FILE REFERENCE: Lea 35693 Foreign Countries
; CURRENT APPLICATION NUMBER: US/10/219,051B
; CURRENT FILING DATE: 2003-05-09
; PRIOR APPLICATION NUMBER: US 60/312,147
; PRIOR FILING DATE: 2001-08-14
```

```
; PRIOR APPLICATION NUMBER: US 60/346,382
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/333,347
; PRIOR FILING DATE: 2001-11-26
; NUMBER OF SEQ ID NOS: 14715
; SOFTWARE: Perl script
; SEQ ID NO 14342
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: SWISS-Prot / P00480
; DATABASE ENTRY DATE: 2002-06-15
US-10-219-051B-14342

Query Match          60.1%; Score 125; DB 28; Length 354;
Best Local Similarity 66.7%; Pred. No. 8.4e-10;
Matches 28; Conservative 3; Mismatches 9; Indels 2; Gaps 1;

QY 1 MLSNRLILNKAALRKAHTSMVRNFRYKGPVQS--QLKPRDL 40
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MLFNLRLILNNAAFRNGHFMVVRNFRCGQPLQNKVKLGKRD 42
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 19
US-09-791-537-53980
; Sequence 53980, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 53980
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Bos taurus
US-09-791-537-53980

Query Match          54.8%; Score 114; DB 22; Length 354;
Best Local Similarity 61.9%; Pred. No. 4.2e-08;
Matches 26; Conservative 4; Mismatches 10; Indels 2; Gaps 1;

QY 1 MLSNRLILNKAALRKAHTSMVRNFRYKGPVQS--QLKPRDL 40
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MLFHLRLILNNAALRNGHFMVVRNFRCGQPLQDKVKLGKRD 42
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 20
PCT-US02-02814-48
; Sequence 48, Application PC/TUS0202814
; GENERAL INFORMATION:
; APPLICANT: University of Rochester
; APPLICANT: Zauderer, Maurice (U.S. Only)
; APPLICANT: Smith, Ernest S. (U.S. Only)
; TITLE OF INVENTION: Methods of Identifying Regulator Molecules
; FILE REFERENCE: 1821.0080003
; CURRENT APPLICATION NUMBER: PCT/US02/02814
; CURRENT FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: 60/271,423
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/265,880
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/265,589
; PRIOR FILING DATE: 2001-02-02
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48
```







```

; LOCATION: (32)..(32)
; ; OTHER INFORMATION: May be any amino acid
US-10-061-395-48

Query Match 44.2%; Score 92; DB 26; Length 32;
Best Local Similarity 61.3%; Pred. No. 5e-06;
Matches 19; Conservative 2; Mismatches 10; Indels 0;

QY 1 MLSNRLILNKAAALRKAHTSMVRNFRYKPV 31
   ||| ||| ||| ||| ||| ||| ||| |||
Db 1 MLFNLRXXLNNAAFRHGHNFVMVNRFCGQPL 31

RESULT 33
US-10-823-254-23
; Sequence 23, Application US/10823254
; GENERAL INFORMATION:
; APPLICANT: Kiener, Peter
; APPLICANT: Kinch, Michael
; APPLICANT: Langermann, Solomon
; APPLICANT: Reed, Jennifer
; TITLE OF INVENTION: Epha2 and Hyperproliferative Cell Disorders
; FILE REFERENCE: 10271-058-999
; CURRENT APPLICATION NUMBER: US/10/823,254
; CURRENT FILING DATE: 2004-04-12
; PRIOR APPLICATION NUMBER: 60/462,024
; PRIOR FILING DATE: 2003-04-11
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 23
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (7)..(8)
; ; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; ;
; NAME/KEY: misc feature
; LOCATION: (32)..(32)
; ; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-10-823-254-23

Query Match 44.2%; Score 92; DB 34; Length 32;
Best Local Similarity 61.3%; Pred. No. 5e-06;
Matches 19; Conservative 2; Mismatches 10; Indels 0;

QY 1 MLSNRLILNKAAALRKAHTSMVRNFRYKPV 31
   ||| ||| ||| ||| ||| ||| ||| |||
Db 1 MLFNLRXXLNNAAFRHGHNFVMVNRFCGQPL 31

RESULT 34
US-10-823-259-23
; Sequence 23, Application US/10823259
; GENERAL INFORMATION:
; APPLICANT: Kiener, Peter
; APPLICANT: Kinch, Michael
; APPLICANT: Langermann, Solomon
; TITLE OF INVENTION: Epha2 and Hyperproliferative Cell Disorders and
; TITLE OF INVENTION: Endothelial Reconstitution
; FILE REFERENCE: 10271-058-999
; CURRENT APPLICATION NUMBER: US/10/823,259
; CURRENT FILING DATE: 2004-04-12
; PRIOR APPLICATION NUMBER: 60/462,009
; PRIOR FILING DATE: 2003-04-11
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 23
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:

```

```
; NAME/KEY: misc_feature
; LOCATION: (7)..(8)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (32)..(32)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-10-823-259-23

Query Match      44.2%; Score 92; DB 34; Length 32;
Best Local Similarity 61.3%; Pred. No. 5e-06;
Matches 19; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLSNRLILNKAALRKAHTSMVNFYKPKV 31
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MLFNLRXLLNNAAFRHGHFNFWNFRGQPL 31

RESULT 35
US-09-791-537-53995
; Sequence 53995, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 53995
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Trachemys scripta elegans
US-09-791-537-53995

Query Match      44.2%; Score 92; DB 22; Length 354;
Best Local Similarity 52.4%; Pred. No. 0.00011;
Matches 22; Conservative 6; Mismatches 12; Indels 2; Gaps 1;

QY 1 MLSNRLILNKAALRKAHTSMVNFYKPKVQS--QLKPRDL 40
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MLFNLRNLLNNAATLRNSSKQLVQHFRSGQPTQTINILKGRDL 42

RESULT 36
US-09-270-849B-181269
; Sequence 181269, Application US/09270849B
; GENERAL INFORMATION:
; APPLICANT: Swimmer et al.
; APPLICANT: Swimmer et al.
; TITLE OF INVENTION: Insect genome survey devices
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/270,849B
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 195450
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 181269
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: at any location having Xaa, Xaa means any amino acid
US-09-270-849B-181269

Query Match      28.8%; Score 60; DB 16; Length 205;
Best Local Similarity 42.1%; Pred. No. 4.7;
Matches 16; Conservative 5; Mismatches 13; Indels 4; Gaps 1;

QY 8 LLNKAALRKAHTSMVNFYKPKVQSOLKPRDLC 41
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
```

```
Db 66 LINVAAPHKAYTILIRLRLFFSYKSHVKSXKIXFLC 103

RESULT 37
US-09-855-604-639
; Sequence 639, Application US/09855604
; GENERAL INFORMATION:
; APPLICANT: GICQUEL, BRIGITTE
; APPLICANT: PORTNOI, DENIS
; APPLICANT: LIM, ENG-MONG
; APPLICANT: PELICIC, VLADIMIR
; APPLICANT: GUIGUENO, AGNES
; APPLICANT: GOGUET DE LA SALMONIERE, YVES
; TITLE OF INVENTION: POLYPEPTIDE NUCLEIC SEQUENCES EXPORTED FROM MYCOBACTERIA,
; TITLE OF INVENTION: VECTORS COMPRISING SAME AND USES FOR DIAGNOSING AND
; TITLE OF INVENTION: PREVENTING TUBERCULOSIS
; FILE REFERENCE: 03715.0062-01000
; CURRENT APPLICATION NUMBER: US/09/855,604
; CURRENT FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 09/485,536
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: PCT/FR98/01813
; PRIOR FILING DATE: 1998-08-14
; PRIOR APPLICATION NUMBER: FR 97 10404
; PRIOR FILING DATE: 1997-08-14
; PRIOR APPLICATION NUMBER: FR 97 11325
; PRIOR FILING DATE: 1997-09-11
; NUMBER OF SEQ ID NOS: 935
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 639
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-855-604-639

Query Match      27.9%; Score 58; DB 23; Length 99;
Best Local Similarity 39.4%; Pred. No. 3.8;
Matches 13; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

QY 6 RILLNKAALRKAHTSMVNFYKPKVQSOLKPR 38
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Db 53 RVILRCATRKANQSRARTLRPLRLPLRIALRPR 85

RESULT 38
US-09-855-604A-639
; Sequence 639, Application US/09855604A
; GENERAL INFORMATION:
; APPLICANT: GICQUEL, BRIGITTE
; APPLICANT: PORTNOI, DENIS
; APPLICANT: LIM, ENG-MONG
; APPLICANT: PELICIC, VLADIMIR
; APPLICANT: GUIGUENO, AGNES
; APPLICANT: GOGUET DE LA SALMONIERE, YVES
; TITLE OF INVENTION: POLYPEPTIDE NUCLEIC SEQUENCES EXPORTED FROM MYCOBACTERIA,
; TITLE OF INVENTION: VECTORS COMPRISING SAME AND USES FOR DIAGNOSING AND
; TITLE OF INVENTION: PREVENTING TUBERCULOSIS
; FILE REFERENCE: 03715.0062-01000
; CURRENT APPLICATION NUMBER: US/09/855,604A
; CURRENT FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 09/485,536
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: PCT/FR98/01813
; PRIOR FILING DATE: 1998-08-14
; PRIOR APPLICATION NUMBER: FR 97 10404
; PRIOR FILING DATE: 1997-08-14
; PRIOR APPLICATION NUMBER: FR 97 11325
; PRIOR FILING DATE: 1997-09-11
; NUMBER OF SEQ ID NOS: 935
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 639
; LENGTH: 99
; TYPE: PRT
```

Search completed: December 18, 2004, 03:03:04  
Job time : 325.095 secs

```
; ORGANISM: Mycobacterium tuberculosis
US-09-855-604A-639

Query Match      27.9%; Score 58; DB 23; Length 99;
Best Local Similarity 39.4%; Pred. No. 3.8;
Matches 13; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

QY 6 RLLNKAALRKAHTSMVRNFRYKGPVQSOLKPR 38
DB 53 RVILRCAHTRKANQGRARTRLPLRLIALRPR 85

RESULT 39
US-09-791-537-85254
; Sequence 85254, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 85254
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Rana catesbeiana
US-09-791-537-85254

Query Match      27.4%; Score 57; DB 22; Length 350;
Best Local Similarity 40.0%; Pred. No. 27;
Matches 16; Conservative 9; Mismatches 13; Indels 2; Gaps 2;

QY 1 MLSNRLILINKAALRKAHTSMVRNFRYKGPVQSOLKPRDL 40
DB 1 MLHMRITIN-ASWRYGNKCIVRQGFQSQ-TYSQLKGRDL 38

RESULT 40
US-10-369-493-3761
; Sequence 3761, Application US/10369493
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 3761
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Neurospora crassa
US-10-369-493-3761

Query Match      26.9%; Score 56; DB 29; Length 302;
Best Local Similarity 64.7%; Pred. No. 32;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 12 AALRKAHTSMVRNFRYK 28
DB 271 AAMTKAHVSQMRWFRYK 287
```

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 18, 2004, 02:45:56 ; Search time 15.131 Seconds  
(without alignments)  
122.155 Million cell updates/sec

Title: US-08-765-244-1  
Perfect score: 208  
Sequence: 1 MLSNRLILNKAAALRKAHTS.....VNFYRGKPVQSLKPRDL 41

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 168898 seqs, 45081143 residues

Total number of hits satisfying chosen parameters: 168898

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending Patents AA New.\*  
1: /cgn2\_6/prodata/2/paa/pct\_NEW\_COMB.pep.\*  
2: /cgn2\_6/prodata/2/paa/US06\_NEW\_COMB.pep.\*  
3: /cgn2\_6/prodata/2/paa/US07\_NEW\_COMB.pep.\*  
4: /cgn2\_6/prodata/2/paa/US08\_NEW\_COMB.pep.\*  
5: /cgn2\_6/prodata/2/paa/US09\_NEW\_COMB.pep.\*  
6: /cgn2\_6/prodata/2/paa/US10\_NEW\_COMB.pep.\*  
7: /cgn2\_6/prodata/2/paa/US11\_NEW\_COMB.pep.\*  
8: /cgn2\_6/prodata/2/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	144	69.2	366	1 PCT-US04-35137-101	Sequence 101, App
2	144	69.2	366	6 US-10-972-963-101	Sequence 101, App
3	125	60.1	354	1 PCT-US04-35137-100	Sequence 100, App
4	125	60.1	354	6 US-10-972-963-100	Sequence 100, App
5	125	60.1	354	6 US-10-990-328-9578	Sequence 9578, App
6	125	60.1	354	6 US-10-990-328-9579	Sequence 9579, App
7	53	25.5	435	6 US-10-732-923-2547	Sequence 2547, App
8	51.5	24.8	200	1 PCT-US02-09107B-57678	Sequence 57678, A
9	50	24.0	882	6 US-10-982-512-1	Sequence 1, Appli
10	49.5	23.8	1542	6 US-10-732-923-1751	Sequence 1751, Ap
11	49	23.6	201	1 PCT-US02-09107B-55605	Sequence 55605, A
12	49	23.6	523	1 PCT-US02-09107B-73109	Sequence 73109, A
13	49	23.6	531	1 PCT-US02-09107B-56536	Sequence 56536, A
14	49	23.6	531	1 PCT-US02-09107B-59504	Sequence 59504, A
15	49	23.6	531	1 PCT-US02-09107B-75990	Sequence 75990, A
16	49	23.6	1903	1 PCT-US04-02460-3	Sequence 3, Appli
17	48	23.1	731	1 PCT-US02-09107B-49817	Sequence 49817, A
18	47.5	22.8	135	8 US-60-622-712-1944	Sequence 1944, Ap
19	47.5	22.8	294	1 PCT-US02-09107B-68798	Sequence 68798, A
20	47.5	22.8	417	6 US-10-965-898-67	Sequence 67, Appl
21	47	22.6	203	6 US-10-931-081A-174	Sequence 174, App
22	47	22.6	310	6 US-10-777-288A-3661	Sequence 3661, App
23	47	22.6	328	1 PCT-US03-29415-127	Sequence 127, App
24	47	22.6	367	1 PCT-US02-09107B-49563	Sequence 49563, A
25	47	22.6	348	6 US-10-990-328-12948	Sequence 12948, A

ALIGNMENTS

RESULT 1  
PCT-US04-35137-101  
; Sequence 101, Application PC/TUS0435137  
; GENERAL INFORMATION:  
; APPLICANT: Gencia Corporation  
; APPLICANT: Khan, Shaharyar  
; TITLE OF INVENTION: Methods and Compositions for the Introduction of Polynucleotides  
; FILE REFERENCE: 120701-2030  
; CURRENT APPLICATION NUMBER: PCT/US04/35137  
; CURRENT FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: 60/568,436  
; PRIOR FILING DATE: 2004-05-05  
; PRIOR APPLICATION NUMBER: 60/513,983  
; PRIOR FILING DATE: 2003-10-24  
; NUMBER OF SEQ ID NOS: 218  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 101  
; LENGTH: 366  
; TYPE: PRT  
; ORGANISM: Mus musculus  
PCT-US04-35137-101

Query Match 69.2%; Score 144; DB 1; Length 366;  
Best Local Similarity 78.6%; Pred. No. 5e-13; Mismatches 2; Indels 2; Gaps 1;  
Matches 33; Conservative 5; Indels 2; Gaps 1;  
QY 1 MLSNRLILNKAAALRKAHTSMVNFYRGKPVQS--QLKPRDL 40  
|||||  
Db 1 MLSNRLILNKAAALRKAHTSVVRFWCGKPVQSQVQLKGRDL 42  
|||||  
RESULT 2  
US-10-972-963-101  
; Sequence 101, Application US/10972963  
; GENERAL INFORMATION:  
; APPLICANT: Gencia Corporation  
; APPLICANT: Khan, Shaharyar  
; TITLE OF INVENTION: Methods and Compositions for the Introduction of Polynucleotides  
; FILE REFERENCE: 120701-2030  
; CURRENT APPLICATION NUMBER: US/10/972,963  
; CURRENT FILING DATE: 2004-10-25  
; PRIOR APPLICATION NUMBER: 60/568,436  
; PRIOR FILING DATE: 2004-05-05  
; PRIOR APPLICATION NUMBER: 60/513,983  
; PRIOR FILING DATE: 2003-10-24  
; NUMBER OF SEQ ID NOS: 218  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 101

```
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-972-963-101

Query Match      69.2%; Score 144; DB 6; Length 366;
Best Local Similarity 78.6%; Pred. No. 5e-13;
Matches 33; Conservative 2; Mismatches 5; Indels 2; Gaps 1;

QY 1 MLSNRLILNKAALRKAHTSMVNFYRGKPVQS--OLKPRDL 40
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MLSNRLILNKAALRKAHTSMVNFYRGKPVQS--OLKPRDL 40
    ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 3
PCT-US04-35137-100
; Sequence 100, Application PC/TUS0435137
; GENERAL INFORMATION:
; APPLICANT: Khan, Shaharyar
; TITLE OF INVENTION: Methods and Compositions for the Introduction of Polynucleotides
; FILE REFERENCE: 120701-2030
; CURRENT APPLICATION NUMBER: PCT/US04/35137
; CURRENT FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: 60/568,436
; PRIOR FILING DATE: 2004-05-05
; PRIOR APPLICATION NUMBER: 60/513,983
; PRIOR FILING DATE: 2003-10-24
; NUMBER OF SEQ ID NOS: 218
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 100
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US04-35137-100

Query Match      60.1%; Score 125; DB 1; Length 354;
Best Local Similarity 66.7%; Pred. No. 2.6e-10;
Matches 28; Conservative 3; Mismatches 9; Indels 2; Gaps 1;

QY 1 MLSNRLILNKAALRKAHTSMVNFYRGKPVQS--OLKPRDL 40
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MLSNRLILNKAALRKAHTSMVNFYRGKPVQS--OLKPRDL 40
    ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 4
US-10-972-963-100
; Sequence 100, Application US/10972963
; GENERAL INFORMATION:
; APPLICANT: Gencia Corporation
; TITLE OF INVENTION: Methods and Compositions for the Introduction of Polynucleotides
; FILE REFERENCE: 120701-2030
; CURRENT APPLICATION NUMBER: US/10/972,963
; CURRENT FILING DATE: 2004-10-25
; PRIOR APPLICATION NUMBER: 60/568,436
; PRIOR FILING DATE: 2004-05-05
; PRIOR APPLICATION NUMBER: 60/513,983
; PRIOR FILING DATE: 2003-10-24
; NUMBER OF SEQ ID NOS: 218
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 100
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-972-963-100

Query Match      60.1%; Score 125; DB 6; Length 354;
Best Local Similarity 66.7%; Pred. No. 2.6e-10;
Matches 28; Conservative 3; Mismatches 9; Indels 2; Gaps 1;

QY 1 MLSNRLILNKAALRKAHTSMVNFYRGKPVQS--OLKPRDL 40
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MLSNRLILNKAALRKAHTSMVNFYRGKPVQS--OLKPRDL 40
    ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 5
US-10-990-328-9578
; Sequence 9578, Application US/10990328
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; FILE REFERENCE: CL001495
; CURRENT APPLICATION NUMBER: US/10/990,328
; CURRENT FILING DATE: 2004-11-17
; NUMBER OF SEQ ID NOS: 558824
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9578
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-990-328-9578

Query Match      60.1%; Score 125; DB 6; Length 354;
Best Local Similarity 66.7%; Pred. No. 2.6e-10;
Matches 28; Conservative 3; Mismatches 9; Indels 2; Gaps 1;

QY 1 MLSNRLILNKAALRKAHTSMVNFYRGKPVQS--OLKPRDL 40
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MLSNRLILNKAALRKAHTSMVNFYRGKPVQS--OLKPRDL 40
    ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 6
US-10-990-328-9579
; Sequence 9579, Application US/10990328
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; FILE REFERENCE: CL001495
; CURRENT APPLICATION NUMBER: US/10/990,328
; CURRENT FILING DATE: 2004-11-17
; NUMBER OF SEQ ID NOS: 558824
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9579
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-990-328-9579

Query Match      60.1%; Score 125; DB 6; Length 354;
Best Local Similarity 66.7%; Pred. No. 2.6e-10;
Matches 28; Conservative 3; Mismatches 9; Indels 2; Gaps 1;

QY 1 MLSNRLILNKAALRKAHTSMVNFYRGKPVQS--OLKPRDL 40
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MLSNRLILNKAALRKAHTSMVNFYRGKPVQS--OLKPRDL 40
    ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 7
US-10-732-923-2547
; Sequence 2547, Application US/10732923
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 2547
; LENGTH: 435
```

RESULT 12

```

; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/342,923
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: 10/072,851
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/362,699
; PRIOR FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 56536
; LENGTH: 531
; TYPE: PRT
; ORGANISM: Escherichia coli
PCT-US02-09107B-56536

Query Match      23.6%; Score 49; DB 1; Length 531;
Best Local Similarity 48.0%; Pred. No. 35;
Matches 12; Conservative 2; Mismatches 7; Indels 4; Gaps 1;

QY      16 KAHTSMVRNFRYKGPVQSOLKPRDL 40
Db      352 KYHT----GFRHAMPVRQLKTRTL 372

RESULT 14
PCT-US02-09107B-59504
; Sequence 59504, Application PC/TUS0209107B
; GENERAL INFORMATION:
; APPLICANT: Elittra Pharmaceuticals Inc.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034VPC
; CURRENT APPLICATION NUMBER: PCT/US02/09107B
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/342,923
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: 10/072,851
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/362,699
; PRIOR FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 59504
; LENGTH: 531
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
PCT-US02-09107B-59504

Query Match      23.6%; Score 49; DB 1; Length 531;
Best Local Similarity 48.0%; Pred. No. 35;
Matches 12; Conservative 2; Mismatches 7; Indels 4; Gaps 1;

QY      16 KAHTSMVRNFRYKGPVQSOLKPRDL 40
Db      352 KYHT----GFRHAMPVRQLKTRTL 372

RESULT 15
PCT-US02-09107B-75990
; Sequence 75990, Application PC/TUS0209107B
; GENERAL INFORMATION:
; APPLICANT: Elittra Pharmaceuticals Inc.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034VPC
; CURRENT APPLICATION NUMBER: PCT/US02/09107B
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/815,242
; PRIOR FILING DATE: 2001-03-21

```

```
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/342,923
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: 10/072,851
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/362,699
; PRIOR FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 75990
; LENGTH: 531
; TYPE: PRT
; ORGANISM: Salmonella typhi
PCT-US02-09107B-75990

Query Match      23.6%; Score 49; DB 1; Length 531;
Best Local Similarity 48.0%; Pred. No. 35;
Matches 12; Conservative 2; Mismatches 7; Indels 4; Gaps 1;

QY      16 KAHTSMVRNFRYKPVQSQKPRDL 40
Db      352 KYHT----GFRHAMPVRQQLKTRTL 372

RESULT 16
PCT-US04-02460-3
; Sequence 3, Application PC/TUS0402460
; GENERAL INFORMATION:
; APPLICANT: Chang, Chia-Hwa
; APPLICANT: Liu, Xiaowen
; APPLICANT: Lewicki, John A.
; APPLICANT: Xu, Qiang
; APPLICANT: Osel, Inc.
; TITLE OF INVENTION: Surface Expression of Biologically Active Proteins in
; FILE REFERENCE: 016976-000810PC
; CURRENT APPLICATION NUMBER: PCT/US04/02460
; CURRENT FILING DATE: 2004-01-28
; PRIOR APPLICATION NUMBER: US 60/443,619
; PRIOR FILING DATE: 2003-01-29
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1903
; TYPE: PRT
; ORGANISM: Lactobacillus jensenii 1153
; FEATURE:
; OTHER INFORMATION: genomic C370 cell wall anchor sequence
; NAME/KEY: DOMAIN
; LOCATION: (1309)..(1903)
; OTHER INFORMATION: CWA200 cell wall associated region with anchor
; OTHER INFORMATION: motif
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1868)..(1872)
; OTHER INFORMATION: anchor motif, sorting signal, cell wall targeting
; OTHER INFORMATION: region
PCT-US04-02460-3

Query Match      23.6%; Score 49; DB 1; Length 1903;
Best Local Similarity 40.0%; Pred. No. 1.4e+02;
Matches 14; Conservative 3; Mismatches 16; Indels 2; Gaps 1;

QY      6 RIILNKAALRKATSMVRNFRYKPVQ--SQLKPR 38
Db      1186 RAALNGDAREKARQELNNYNSGKSLQDGSITLDP 1220

RESULT 17
PCT-US02-09107B-49817
; Sequence 49817, Application PC/TUS0209107B
; GENERAL INFORMATION:
; APPLICANT: Elittra Pharmaceuticals Inc.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034VPC
; CURRENT APPLICATION NUMBER: PCT/US02/09107B
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/362,699
; PRIOR FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 49817
; LENGTH: 731
; TYPE: PRT
; ORGANISM: Burkholderia fungorum
PCT-US02-09107B-49817

Query Match      23.1%; Score 48; DB 1; Length 731;
Best Local Similarity 45.5%; Pred. No. 69;
Matches 10; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY      4 NLRILNKAALRKATSMVRNF 25
Db      559 NLRINVSARQLVDTAVRHF 580

RESULT 18
US-60-622-712-1944
; Sequence 1944, Application US/60622712
; GENERAL INFORMATION:
; APPLICANT: Klaenhammer, Todd
; APPLICANT: Russell, William
; APPLICANT: Alterman, Eric
; TITLE OF INVENTION: LACTOBACILLUS ACIDOPHILUS NUCLEIC ACIDS AND USES THEREOF
; FILE REFERENCE: 5051-604PR10
; CURRENT APPLICATION NUMBER: US/60/622,712
; CURRENT FILING DATE: 2004-10-27
; NUMBER OF SEQ ID NOS: 2559
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 1944
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Lactobacillus acidophilus
US-60-622-712-1944

Query Match      22.8%; Score 47.5; DB 8; Length 135;
Best Local Similarity 27.3%; Pred. No. 13;
Matches 12; Conservative 9; Mismatches 12; Indels 11; Gaps 1;

QY      4 NLRILNKAALRKATSMVRNFR-----YKPVQSQK 36
Db      50 NSCIVMSAGMITKAHQLDVKFNFNKNYAKTKRMGYGKSVVTEMR 93

RESULT 19
PCT-US02-09107B-68798
; Sequence 68798, Application PC/TUS0209107B
; GENERAL INFORMATION:
; APPLICANT: Elittra Pharmaceuticals Inc.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034VPC
; CURRENT APPLICATION NUMBER: PCT/US02/09107B
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 09/948,993
```

```

;
; CLONE: 3235839
; SEQUENCE DESCRIPTION: SEQ ID NO: 67 :
US-10-965-898-67

Query Match                22.8%; Score 47.5; DB 6; Length 417;
Best Local Similarity      37.8%; Pred. No. 45;
Matches 14; Conservative 3; Mismatches 19; Indels 1; Gaps 1;

QY 5 LRLLNKAALRKAHTSMVRNFRYRKPVQSQLKPRDLC 41
   |||:|||:|||:|||:|||:|||:|||:|||:|||
Db 17 LRSLINAAQLTKFTRTPARTLLHGFSAPQIS-SDNC 52

RESULT 21
US-10-931-081A-174
; Sequence 174, Application US/10931081A
; GENERAL INFORMATION:
; APPLICANT: Lough, Tony James
; APPLICANT: Hermameier, Dieter H.
; APPLICANT: Varkonyi-Gasic, Erika
; APPLICANT: Sweetman, Justin
; APPLICANT: Havukkala, Ilkka J
; APPLICANT: Belanger, Helene
; APPLICANT: Forster, Richard L.S.
; APPLICANT: Hudson, Keith R.
; TITLE OF INVENTION: Control of Floral Induction
; FILE REFERENCE: 1084U
; CURRENT APPLICATION NUMBER: US/10/931,081A
; CURRENT FILING DATE: 2004-08-30
; NUMBER OF SEQ ID NOS: 872
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 174
; LENGTH: 203
; TYPE: PRT
; ORGANISM: Cucurbita maxima
US-10-931-081A-174

Query Match                22.6%; Score 47; DB 6; Length 203;
Best Local Similarity      34.8%; Pred. No. 24;
Matches 16; Conservative 4; Mismatches 8; Indels 18; Gaps 2;

QY 12 AALRKAHTSMVR-----NFRYG--KPVQSQLKPRD 39
   |||:|||:|||:|||:|||:|||:|||:|||:|||
Db 30 AALTKAHTSLRLVLTPELQSKVAALVDPNFDGSEKPKRGKKDD 75

RESULT 22
US-10-777-288A-3661
; Sequence 3661, Application US/10777288A
; GENERAL INFORMATION:
; APPLICANT: Research Association for Biotechnology
; TITLE OF INVENTION: Full length cDNA
; FILE REFERENCE: BTR-A0301Y1
; CURRENT APPLICATION NUMBER: US/10/777,288A
; CURRENT FILING DATE: 2004-02-13
; PRIOR APPLICATION NUMBER: JP 2003-102207
; PRIOR FILING DATE: 2003-02-14
; PRIOR APPLICATION NUMBER: JP 2003-131452
; PRIOR FILING DATE: 2003-05-09
; NUMBER OF SEQ ID NOS: 394
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 3661
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-777-288A-3661

Query Match                22.6%; Score 47; DB 6; Length 310;
Best Local Similarity      26.7%; Pred. No. 38;
Matches 8; Conservative 11; Mismatches 9; Indels 2; Gaps 1;

QY 10 NKAALRKAHTSMVRNFR--YGRKPVQSQLKPK 37
   |||:|||:|||:|||:|||:|||:|||:|||:|||

```

RESULT 25  
US-10-990-328-12948  
; Sequence 12948, Application US/10990328  
; GENERAL INFORMATION:

```
; PRIOR APPLICATION NUMBER: 60/342,923
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: 10/072,851
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/362,699
; PRIOR FILING DATE: 2002-03-06
; PRIOR FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 70268
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
PCT-US02-09107B-70268

Query Match      22.4%; Score 46.5; DB 1; Length 351;
Best Local Similarity 25.6%; Pred. No. 52;
Matches 11; Conservative 8; Mismatches 11; Indels 13; Gaps 1;

QY 8 LNKAAALRKAHT-----SMVRNFRYKGPVQSOLKP 37
   ||| |||
Db 30 LVDKPNRKVHTKPISVMGTVILFSLGIGWIGHPIETBIKP 72

RESULT 28
PCT-US02-09107B-58201
; Sequence 58201, Application PC/TUS0209107B
; GENERAL INFORMATION:
; APPLICANT: Elittra Pharmaceuticals Inc.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034VPC
; CURRENT APPLICATION NUMBER: PCT/US02/09107B
; CURRENT FILING DATE: 2002-03-12
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/342,923
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: 10/072,851
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/362,699
; PRIOR FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 58201
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
PCT-US02-09107B-58201

Query Match      22.4%; Score 46.5; DB 1; Length 386;
Best Local Similarity 37.9%; Pred. No. 57;
Matches 11; Conservative 4; Mismatches 13; Indels 1; Gaps 1;

QY 11 KAALRKAHTSMVRNFRYKGPVQSOLKPRD 39
   ||| |||
Db 45 KLALEKGAQVMVTS-HLGRPTGEFEKPED 72

RESULT 29
PCT-US02-09107B-46555
; Sequence 46555, Application PC/TUS0209107B
; GENERAL INFORMATION:
; APPLICANT: Elittra Pharmaceuticals Inc.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034VPC
; CURRENT APPLICATION NUMBER: PCT/US02/09107B
; CURRENT FILING DATE: 2002-03-12
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/342,923
```

```
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: 10/072,851
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/362,699
; PRIOR FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 46555
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Bacillus anthracis
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (14)..(14)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (189)..(189)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (202)..(202)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (254)..(254)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (268)..(268)
; OTHER INFORMATION: X=any amino acid
PCT-US02-09107B-46555

Query Match      22.4%; Score 46.5; DB 1; Length 510;
Best Local Similarity 34.4%; Pred. No. 77;
Matches 11; Conservative 6; Mismatches 6; Indels 9; Gaps 1;

QY 12 AALRKAHT-----SMVRNFRYKGPVQSOLKP 34
   ||| |||
Db 261 ADIRGAHXTGQKVNKIALVRNVREGEPINSR 292

RESULT 30
US-10-952-698-43
; Sequence 43, Application US/10952698
; GENERAL INFORMATION:
; APPLICANT: Mack, David
; APPLICANT: Aziz, Nataeha
; TITLE OF INVENTION: METHODS OF DIAGNOSIS OF BLADDER CANCER, COMPOSITIONS AND METHODS
; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: US/10/952,698
; CURRENT FILING DATE: 2004-09-28
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 43
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-952-698-43

Query Match      22.4%; Score 46.5; DB 6; Length 513;
Best Local Similarity 33.3%; Pred. No. 78;
Matches 12; Conservative 11; Mismatches 12; Indels 1; Gaps 1;

QY 1 MLSNRLILANKAALRKAHTSMVR-NERYKGPVQSOL 35
   ||:|||||:|||||:|||||:|||||:|
Db 66 LLASLIQILWKGKGRKQHDITLVEYHKYKIFRMKL 101

RESULT 31
US-10-220-366A-20931
; Sequence 20931, Application US/10220366A
```



```
; GENERAL INFORMATION:
; APPLICANT: HISEQ, INC
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-042
; CURRENT APPLICATION NUMBER: US/10/220,366A
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 09/577,409
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: 09/515,126
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 27802
; SOFTWARE: Custom
; SEQ ID NO 20931
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-220-366A-20931
```

```
Query Match 22.1%; Score 46; DB 6; Length 98;
Best Local Similarity 34.4%; Pred. No. 15;
Matches 11; Conservative 7; Mismatches 12; Indels 2; Gaps 1;

QY 6 RILLKAALRKHAHTSMVRNFRYG--KPVOSQL 35
|::||| |::| |::| |::| |::|
Db 2 RLVINSELARSALSLPKCWDYGPEPPCPAQM 33
```

```
RESULT 32
PCT-US02-09107B-77890
; Sequence 77890, Application PC/TUS0209107B
; GENERAL INFORMATION:
; APPLICANT: Elittra Pharmaceuticals Inc.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034VPC
; CURRENT APPLICATION NUMBER: PCT/US02/09107B
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/342,923
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: 10/072,851
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/362,699
; PRIOR FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 77890
; LENGTH: 333
; TYPE: PRT
; ORGANISM: Yersinia pestis
PCT-US02-09107B-77890
```

```
Query Match 22.1%; Score 46; DB 1; Length 333;
Best Local Similarity 47.8%; Pred. No. 57;
Matches 11; Conservative 2; Mismatches 6; Indels 4; Gaps 1;

QY 18 HTSMVRNFRYKPVOSQLKPRDL 40
||| ||| ||| ||| ||| |||
Db 156 HT----GFRHMPVRQOLKTRTL 174
```

```
RESULT 33
PCT-US02-09107B-51634
; Sequence 51634, Application PC/TUS0209107B
; GENERAL INFORMATION:
; APPLICANT: Elittra Pharmaceuticals Inc.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034VPC
; CURRENT APPLICATION NUMBER: PCT/US02/09107B
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/815,242
```

```
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/342,923
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: 10/072,851
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/362,699
; PRIOR FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51634
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Clostridium acetobutylicum
PCT-US02-09107B-51634
```

```
Query Match 22.1%; Score 46; DB 1; Length 351;
Best Local Similarity 29.4%; Pred. No. 61;
Matches 10; Conservative 7; Mismatches 17; Indels 0; Gaps 0;

QY 3 SNLRILLNKAALRKHAHTSMVRNFRYKPVOSQLK 36
|::||| |::| |::| |::| |::|
Db 50 SGVSVVIQKATQGLYHNDLSLLNRYPKITNAGLK 83
```

```
RESULT 34
PCT-US02-09107B-61451
; Sequence 61451, Application PC/TUS0209107B
; GENERAL INFORMATION:
; APPLICANT: Elittra Pharmaceuticals Inc.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034VPC
; CURRENT APPLICATION NUMBER: PCT/US02/09107B
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/342,923
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: 10/072,851
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/362,699
; PRIOR FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 61451
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Legionella pneumophila
PCT-US02-09107B-61451
```

```
Query Match 22.1%; Score 46; DB 1; Length 513;
Best Local Similarity 26.1%; Pred. No. 92;
Matches 12; Conservative 12; Mismatches 10; Indels 12; Gaps 2;

QY 1 MLNLRILLNKAALRKHAHTSMVRNFRYG-----KPVOSQLKPR 38
|::| |::| |::| |::| |::| |::|
Db 151 LLSHGQIILNS---KGHSTVLKNIQIGAEQFNLKNPSFVSQIKAK 192
```

```
RESULT 35
PCT-US02-09107B-77442
; Sequence 77442, Application PC/TUS0209107B
; GENERAL INFORMATION:
; APPLICANT: Elittra Pharmaceuticals Inc.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034VPC
; CURRENT APPLICATION NUMBER: PCT/US02/09107B
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/815,242
; PRIOR FILING DATE: 2001-03-21
```

```
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/342,923
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: 10/072,851
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/362,699
; PRIOR FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 77442
; LENGTH: 1487
; TYPE: PRT
; ORGANISM: Vibrio cholerae
PCT-US02-09107B-77442

Query Match      22.1%; Score 46; DB 1; Length 1487;
Best Local Similarity 28.6%; Pred. No. 2.9e+02;
Matches 6; Conservative 11; Mismatches 4; Indels 0; Gaps 0;

QY      8 LLNKAALRKHAHTSMVRNFRYTG 28
Db      686 MIDDGALQKSYREWMQNYQG 706

RESULT 36
PCT-US02-09107B-44710
; Sequence 44710, Application PC/TUS0209107B
; GENERAL INFORMATION:
; APPLICANT: Elittra Pharmaceuticals Inc.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034VPC
; CURRENT APPLICATION NUMBER: PCT/US02/09107B
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/342,923
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: 10/072,851
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/362,699
; PRIOR FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44710
; LENGTH: 395
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
PCT-US02-09107B-44710

Query Match      21.9%; Score 45.5; DB 1; Length 395;
Best Local Similarity 44.4%; Pred. No. 82;
Matches 12; Conservative 4; Mismatches 10; Indels 1; Gaps 1;

QY      11 KAALRKHAHTSMVRNFRYKGPVQSOLKP 37
Db      45 KAALKGAAVMVFS-HLGRPVEGEPKP 70

RESULT 37
US-10-513-302-2
; Sequence 2, Application US/10513302
; GENERAL INFORMATION:
; APPLICANT: ZHANG, Angi
; APPLICANT: GAO, Mian
; TITLE OF INVENTION: Stable Cytochrome P450 24 (CYP24) Expressing Cell Line and
; FILE REFERENCE: 09539-0008U1
; CURRENT APPLICATION NUMBER: US/10/513,302
; CURRENT FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/376,806

; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/342,923
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: 10/072,851
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/362,699
; PRIOR FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 77442
; LENGTH: 1487
; TYPE: PRT
; ORGANISM: Vibrio cholerae
PCT-US02-09107B-77442

Query Match      22.1%; Score 46; DB 1; Length 1487;
Best Local Similarity 28.6%; Pred. No. 2.9e+02;
Matches 6; Conservative 11; Mismatches 4; Indels 0; Gaps 0;

QY      8 LLNKAALRKHAHTSMVRNFRYTG 28
Db      686 MIDDGALQKSYREWMQNYQG 706

RESULT 36
PCT-US02-09107B-44710
; Sequence 44710, Application PC/TUS0209107B
; GENERAL INFORMATION:
; APPLICANT: Elittra Pharmaceuticals Inc.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034VPC
; CURRENT APPLICATION NUMBER: PCT/US02/09107B
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/342,923
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: 10/072,851
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/362,699
; PRIOR FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44710
; LENGTH: 395
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
PCT-US02-09107B-44710

Query Match      21.9%; Score 45.5; DB 1; Length 395;
Best Local Similarity 44.4%; Pred. No. 82;
Matches 12; Conservative 4; Mismatches 10; Indels 1; Gaps 1;

QY      11 KAALRKHAHTSMVRNFRYKGPVQSOLKP 37
Db      45 KAALKGAAVMVFS-HLGRPVEGEPKP 70

RESULT 37
US-10-513-302-2
; Sequence 2, Application US/10513302
; GENERAL INFORMATION:
; APPLICANT: ZHANG, Angi
; APPLICANT: GAO, Mian
; TITLE OF INVENTION: Stable Cytochrome P450 24 (CYP24) Expressing Cell Line and
; FILE REFERENCE: 09539-0008U1
; CURRENT APPLICATION NUMBER: US/10/513,302
; CURRENT FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/376,806

; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/342,923
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: 10/072,851
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/362,699
; PRIOR FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 77442
; LENGTH: 1487
; TYPE: PRT
; ORGANISM: Vibrio cholerae
PCT-US02-09107B-77442

Query Match      21.9%; Score 45.5; DB 6; Length 514;
Best Local Similarity 33.3%; Pred. No. 1.1e+02;
Matches 12; Conservative 10; Mismatches 13; Indels 1; Gaps 1;

QY      1 MLNLRILLNKAALRKHAHTSMVR-NFRYKGPVQSOL 35
Db      66 LLGSLQLLWKGLKKQHDITLVEYHKYKGIKPRMKL 101

RESULT 38
US-10-990-328-13661
; Sequence 13661, Application US/10990328
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
; FILE REFERENCE: CL001495
; CURRENT APPLICATION NUMBER: US/10/990,328
; CURRENT FILING DATE: 2004-11-17
; NUMBER OF SEQ ID NOS: 558824
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13661
; LENGTH: 514
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-990-328-13661

Query Match      21.9%; Score 45.5; DB 6; Length 514;
Best Local Similarity 33.3%; Pred. No. 1.1e+02;
Matches 12; Conservative 10; Mismatches 13; Indels 1; Gaps 1;

QY      1 MLNLRILLNKAALRKHAHTSMVR-NFRYKGPVQSOL 35
Db      66 LLGSLQLLWKGLKKQHDITLVEYHKYKGIKPRMKL 101

RESULT 39
US-10-408-765-2284
; Sequence 2284, Application US/10408765
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3025
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2284
; LENGTH: 514
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765-2284

Query Match      21.9%; Score 45.5; DB 6; Length 514;
Best Local Similarity 33.3%; Pred. No. 1.1e+02;
Matches 12; Conservative 10; Mismatches 13; Indels 1; Gaps 1;
```

QY 1 MLSNRLILNKALRKAHTSMVR-NFRYKGPVQSOL 35  
Db 66 LGSLLQILWKGGLKKQHDTLVEYHKYKIFRMKL 101

RESULT 40

US-10-732-923-8629  
; Sequence 8629, Application US/10732923  
; GENERAL INFORMATION:  
; APPLICANT: Edgerton, Michael D  
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES  
; FILE REFERENCE: 38-15(52796)C  
; CURRENT APPLICATION NUMBER: US/10/732,923  
; CURRENT FILING DATE: 2003-12-10  
; PRIOR APPLICATION NUMBER: 10/310,154  
; PRIOR FILING DATE: 2002-12-04  
; NUMBER OF SEQ ID NOS: 24149  
; SEQ ID NO 8629  
; LENGTH: 1619  
; TYPE: PRT  
; ORGANISM: Saccharomyces cerevisiae  
US-10-732-923-8629

Query Match 21.9%; Score 45.5; DB 6; Length 1619;  
Best Local Similarity 28.0%; Pred.No. 3.8e+02;  
Matches 14; Conservative 9; Mismatches 12; Indels 15; Gaps 2;

QY 2 LSNRLILNKALRKAHTSMVRNFRYKGPV-----QSOLKPRDL 40  
Db 1190 LRKRVLLNAILRRSKADKID---GKPLLEPPKIVEVDESRLKGEEL 1235

Search completed: December 18, 2004, 03:03:44  
Job time : 17.131 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 18, 2004, 02:41:57 ; Search time 339.905 Seconds  
(without alignments)  
140.169 Million cell updates/sec

Title: US-08-765-244-22

Perfect score: 217

Sequence: 1 MLSNRIILNKRAALRKHTS.....NFRYKPVQSVOLKPRDLIC 43

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 6730630 seqs, 1107998698 residues

Total number of hits satisfying chosen parameters: 6730630

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Pending Patents AA Main:\*
- 1: /cgn2\_6/ptodata/1/paa/PCTUS COMB.pep.\*
  - 2: /cgn2\_6/ptodata/1/paa/US06 COMB.pep.\*
  - 3: /cgn2\_6/ptodata/1/paa/US07 COMB.pep.\*
  - 4: /cgn2\_6/ptodata/1/paa/US08 COMB.pep.\*
  - 5: /cgn2\_6/ptodata/1/paa/US081 COMB.pep.\*
  - 6: /cgn2\_6/ptodata/1/paa/US082 COMB.pep.\*
  - 7: /cgn2\_6/ptodata/1/paa/US083 COMB.pep.\*
  - 8: /cgn2\_6/ptodata/1/paa/US084 COMB.pep.\*
  - 9: /cgn2\_6/ptodata/1/paa/US085 COMB.pep.\*
  - 10: /cgn2\_6/ptodata/1/paa/US086 COMB.pep.\*
  - 11: /cgn2\_6/ptodata/1/paa/US087 COMB.pep.\*
  - 12: /cgn2\_6/ptodata/1/paa/US088 COMB.pep.\*
  - 13: /cgn2\_6/ptodata/1/paa/US089 COMB.pep.\*
  - 14: /cgn2\_6/ptodata/1/paa/US090 COMB.pep.\*
  - 15: /cgn2\_6/ptodata/1/paa/US091 COMB.pep.\*
  - 16: /cgn2\_6/ptodata/1/paa/US092 COMB.pep.\*
  - 17: /cgn2\_6/ptodata/1/paa/US093 COMB.pep.\*
  - 18: /cgn2\_6/ptodata/1/paa/US094 COMB.pep.\*
  - 19: /cgn2\_6/ptodata/1/paa/US095 COMB.pep.\*
  - 20: /cgn2\_6/ptodata/1/paa/US096 COMB.pep.\*
  - 21: /cgn2\_6/ptodata/1/paa/US097A COMB.pep.\*
  - 22: /cgn2\_6/ptodata/1/paa/US097B COMB.pep.\*
  - 23: /cgn2\_6/ptodata/1/paa/US098 COMB.pep.\*
  - 24: /cgn2\_6/ptodata/1/paa/US099A COMB.pep.\*
  - 25: /cgn2\_6/ptodata/1/paa/US099B COMB.pep.\*
  - 26: /cgn2\_6/ptodata/1/paa/US100 COMB.pep.\*
  - 27: /cgn2\_6/ptodata/1/paa/US101 COMB.pep.\*
  - 28: /cgn2\_6/ptodata/1/paa/US102 COMB.pep.\*
  - 29: /cgn2\_6/ptodata/1/paa/US103 COMB.pep.\*
  - 30: /cgn2\_6/ptodata/1/paa/US104 COMB.pep.\*
  - 31: /cgn2\_6/ptodata/1/paa/US105 COMB.pep.\*
  - 32: /cgn2\_6/ptodata/1/paa/US106 COMB.pep.\*
  - 33: /cgn2\_6/ptodata/1/paa/US107 COMB.pep.\*
  - 34: /cgn2\_6/ptodata/1/paa/US108 COMB.pep.\*
  - 35: /cgn2\_6/ptodata/1/paa/US109 COMB.pep.\*
  - 36: /cgn2\_6/ptodata/1/paa/US60 COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

\*

Result No.	Score	Match	Length	DB	ID	Description
1	217	100.0	43	11	US-08-765-244-22	Sequence 22, Appl Sequence 57942, A
2	208	95.9	354	22	US-09-791-537-57942	Sequence 17653, A
3	199	91.7	354	22	US-09-791-537-17653	Sequence 84892, A
4	199	91.7	354	22	US-09-791-537-84892	Sequence 12767, A
5	199	91.7	354	28	US-10-219-051B-12767	Sequence 14340, A
6	199	91.7	354	28	US-10-219-051B-14340	Sequence 1, Appl1
7	197	90.8	41	11	US-08-765-244-1	Sequence 27693, A
8	164	75.6	354	22	US-09-791-537-27693	Sequence 104643, A
9	164	75.6	354	22	US-09-791-537-104643	Sequence 90, Appl
10	164	75.6	366	1	PCT-US04-20454-90	Sequence 89, Appl
11	141	65.0	354	1	PCT-US04-20454-90	Sequence 76239, A
12	141	65.0	354	21	US-09-724-676-76239	Sequence 2694, Ap
13	141	65.0	354	21	US-09-724-676A-76239	Sequence 46324, A
14	141	65.0	354	22	US-09-791-537-46324	Sequence 88687, A
15	141	65.0	354	22	US-09-791-537-88687	Sequence 12769, A
16	141	65.0	354	22	US-09-791-537-12769	Sequence 14342, A
17	141	65.0	354	28	US-10-219-051B-14342	Sequence 53985, A
18	141	65.0	354	28	US-10-219-051B-14342	Sequence 48, Appl
19	130	59.9	354	22	US-09-791-537-53980	Sequence 54, Appl
20	106	48.8	354	22	US-09-791-537-53995	Sequence 23, Appl
21	92	42.4	32	1	PCT-US02-02814-48	Sequence 26, Appl
22	92	42.4	32	1	PCT-US02-21677-54	Sequence 29, Appl
23	92	42.4	32	1	PCT-US04-11481-23	Sequence 29, Appl
24	92	42.4	32	1	PCT-US04-11482-23	Sequence 29, Appl
25	92	42.4	32	1	PCT-US04-23191-26	Sequence 54, Appl
26	92	42.4	32	1	PCT-US04-23192-26	Sequence 48, Appl
27	92	42.4	32	3	US-07-916-939-29	Sequence 23, Appl
28	92	42.4	32	8	US-08-438-190-29	Sequence 23, Appl
29	92	42.4	32	16	US-09-287-145-29	Sequence 29, Appl
30	92	42.4	32	19	US-09-522-727B-29	Sequence 29, Appl
31	92	42.4	32	19	US-09-522-727D-29	Sequence 29, Appl
32	92	42.4	32	26	US-10-052-942-54	Sequence 54, Appl
33	92	42.4	32	26	US-10-061-395-48	Sequence 48, Appl
34	92	42.4	32	34	US-10-823-254-23	Sequence 23, Appl
35	92	42.4	32	34	US-10-823-259-23	Sequence 23, Appl
36	70	32.3	356	22	US-09-791-537-53985	Sequence 53985, A
37	61	28.1	1235	21	US-09-708-427-25599	Sequence 25599, A
38	61	28.1	1247	21	US-09-708-427-25598	Sequence 25598, A
39	61	28.1	1286	1	PCT-US02-18153-3	Sequence 3, Appl1
40	61	28.1	1286	21	US-09-708-427-25597	Sequence 25597, A
41	61	28.1	1286	25	US-09-998-027-3	Sequence 3, Appl1
42	61	28.1	1286	27	US-10-165-099-3	Sequence 3, Appl1
43	58	26.7	627	16	US-09-252-691-8660	Sequence 8660, Ap
44	58	26.7	627	16	US-09-252-691C-8660	Sequence 8660, Ap
45	58	26.7	627	30	US-10-417-886-8660	Sequence 8660, Ap

ALIGNMENTS

RESULT 1  
US-08-765-244-22  
; Sequence 22, Application US/08765244  
; GENERAL INFORMATION:  
; APPLICANT: Seibel, Peter  
; APPLICANT: Seibel, Andrea  
; TITLE OF INVENTION: CHIMERICAL PEPTIDE-NUCLEIC ACID  
; TITLE OF INVENTION: FRAGMENT, PROCESS FOR PRODUCING THE SAME AND ITS USES FOR  
; TITLE OF INVENTION: APPROPRIATELY INTRODUCING NUCLEIC ACIDS INTO CELL ORGANELLES  
; TITLE OF INVENTION: AND CELLS  
; FILE REFERENCE: 8484-0018-999  
; CURRENT APPLICATION NUMBER: US/08/765,244  
; CURRENT FILING DATE: 1997-10-30  
; PRIOR APPLICATION NUMBER: PCT/DE95/00775  
; PRIOR FILING DATE: 1995-06-11  
; PRIOR APPLICATION NUMBER: DE P 44 21 079.5  
; PRIOR FILING DATE: 1994-06-16  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 22

```

; LENGTH: 43
; TYPE: PRT
; ORGANISM: Rattus rattus
US-08-765-244-22

Query Match      100.0%; Score 217; DB 11; Length 43;
Best Local Similarity 100.0%; Pred. No. 2e-25; 0; Indels 0; Gaps 0;
Matches 43; Conservative 0; Mismatches 0;

QY 1 MLSNRLILNKAAALRKAHTSMVRNFRYKPKVQSQVOLKPRDL 43
   |||||
Db 1 MLSNRLILNKAAALRKAHTSMVRNFRYKPKVQSQVOLKPRDL 43
   |||||

RESULT 2
US-09-791-537-57942
; Sequence 57942, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 57942
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-791-537-57942

Query Match      95.9%; Score 208; DB 22; Length 354;
Best Local Similarity 100.0%; Pred. No. 7.7e-23;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLSNRLILNKAAALRKAHTSMVRNFRYKPKVQSQVOLKPRDL 42
   |||||
Db 1 MLSNRLILNKAAALRKAHTSMVRNFRYKPKVQSQVOLKPRDL 42
   |||||

RESULT 3
US-09-791-537-17653
; Sequence 17653, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 17653
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-791-537-17653

Query Match      91.7%; Score 199; DB 22; Length 354;
Best Local Similarity 97.6%; Pred. No. 2e-21;
Matches 41; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLSNRLILNKAAALRKAHTSMVRNFRYKPKVQSQVOLKPRDL 42
   |||||
Db 1 MLSNRLILNKAAALRKAHTSMVRNFRYKPKVQSQVOLKGRDL 42
   |||||

RESULT 4
US-09-791-537-17653
; Sequence 17653, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 17653
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-791-537-17653

Query Match      91.7%; Score 199; DB 22; Length 354;
Best Local Similarity 97.6%; Pred. No. 2e-21;
Matches 41; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLSNRLILNKAAALRKAHTSMVRNFRYKPKVQSQVOLKPRDL 42
   |||||
Db 1 MLSNRLILNKAAALRKAHTSMVRNFRYKPKVQSQVOLKGRDL 42
   |||||

```

```

US-09-791-537-84892
; Sequence 84892, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 84892
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-791-537-84892

Query Match      91.7%; Score 199; DB 22; Length 354;
Best Local Similarity 97.6%; Pred. No. 2e-21;
Matches 41; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLSNRLILNKAAALRKAHTSMVRNFRYKPKVQSQVOLKPRDL 42
   |||||
Db 1 MLSNRLILNKAAALRKAHTSMVRNFRYKPKVQSQVOLKGRDL 42
   |||||

RESULT 5
US-10-219-051B-12767
; Sequence 12767, Application US/10219051B
; GENERAL INFORMATION:
; APPLICANT: The General Hospital Corporation doing business as Massachusetts General
; APPLICANT: Hospital / Bayer AG
; TITLE OF INVENTION: Nucleotide sequences involved in pain
; FILE REFERENCE: Lea 35693 Foreign Countries
; CURRENT APPLICATION NUMBER: US/10/219,051B
; CURRENT FILING DATE: 2003-05-09
; PRIOR APPLICATION NUMBER: US 60/312,147
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/346,382
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/333,347
; PRIOR FILING DATE: 2001-11-26
; NUMBER OF SEQ ID NOS: 14715
; SOFTWARE: Perl script
; SEQ ID NO 12767
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: SWISS-Prot / AAA41767
; DATABASE ENTRY DATE: 2001-10-16
US-10-219-051B-12767

Query Match      91.7%; Score 199; DB 28; Length 354;
Best Local Similarity 97.6%; Pred. No. 2e-21;
Matches 41; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLSNRLILNKAAALRKAHTSMVRNFRYKPKVQSQVOLKPRDL 42
   |||||
Db 1 MLSNRLILNKAAALRKAHTSMVRNFRYKPKVQSQVOLKGRDL 42
   |||||

RESULT 6
US-10-219-051B-14340
; Sequence 14340, Application US/10219051B
; GENERAL INFORMATION:
; APPLICANT: The General Hospital Corporation doing business as Massachusetts General
; APPLICANT: Hospital / Bayer AG
; TITLE OF INVENTION: Nucleotide sequences involved in pain
; FILE REFERENCE: Lea 35693 Foreign Countries
; CURRENT APPLICATION NUMBER: US/10/219,051B

```

; CURRENT FILING DATE: 2003-05-09  
; PRIOR APPLICATION NUMBER: US 60/312,147  
; PRIOR FILING DATE: 2001-08-14  
; PRIOR APPLICATION NUMBER: US 60/346,382  
; PRIOR FILING DATE: 2001-11-01  
; PRIOR APPLICATION NUMBER: US 60/333,347  
; PRIOR FILING DATE: 2001-11-26  
; NUMBER OF SEQ ID NOS: 14715  
; SOFTWARE: Perl script  
; SEQ ID NO 14340  
; LENGTH: 354  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
; PUBLICATION INFORMATION:  
; DATABASE ACCESSION NUMBER: SWISS-Prot / OWRP  
; DATABASE ENTRY DATE: 2001-10-16  
US-10-219-051B-14340

Query Match 91.7%; Score 199; DB 28; Length 354;  
Best Local Similarity 97.6%; Pred. No. 2e-21; Indels 1; Gaps 0;  
Matches 41; Conservative 0; Mismatches 0; Gaps 0;

Qy 1 MSLNRLILNKAAALRKAHTSMVNFYRGKPVQSQVOLKPRDL 42  
Db 1 MSLNRLILNKAAALRKAHTSMVNFYRGKPVQSQVOLKGRDL 42

RESULT 7  
US-08-765-244-1  
; Sequence 1, Application US/08765244  
; GENERAL INFORMATION:  
; APPLICANT: Seibel, Peter  
; APPLICANT: Seibel, Andrea  
; TITLE OF INVENTION: CHEMICAL PEPTIDE-NUCLEIC ACID  
; TITLE OF INVENTION: FRAGMENT, PROCESS FOR PRODUCING THE SAME AND ITS USES FOR  
; TITLE OF INVENTION: APPROPRIATELY INTRODUCING NUCLEIC ACIDS INTO CELL ORGANELLES  
; TITLE OF INVENTION: AND CELLS  
; FILE REFERENCE: 8484-0018-999  
; CURRENT APPLICATION NUMBER: US/08/765,244  
; CURRENT FILING DATE: 1997-10-30  
; PRIOR APPLICATION NUMBER: PCT/DE95/00775  
; PRIOR FILING DATE: 1995-06-11  
; PRIOR APPLICATION NUMBER: DE 44 21 079.5  
; PRIOR FILING DATE: 1994-06-16  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 41  
; TYPE: PRT  
; ORGANISM: Rattus rattus  
; FEATURE:  
US-08-765-244-1

Query Match 90.8%; Score 197; DB 11; Length 41;  
Best Local Similarity 95.3%; Pred. No. 2.4e-22; Indels 2; Gaps 1;  
Matches 41; Conservative 0; Mismatches 0; Gaps 1;

Qy 1 MSLNRLILNKAAALRKAHTSMVNFYRGKPVQSQVOLKPRDLC 43  
Db 1 MSLNRLILNKAAALRKAHTSMVNFYRGKPVQSQVOLKPRDLC 41

RESULT 8  
US-09-791-537-27693  
; Sequence 27693, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
; TITLE OF INVENTION: METHODS OF USE THEREOF  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537

; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 27693  
; LENGTH: 354  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-791-537-27693

Query Match 75.6%; Score 164; DB 22; Length 354;  
Best Local Similarity 83.3%; Pred. No. 5.6e-16;  
Matches 35; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 MSLNRLILNKAAALRKAHTSMVNFYRGKPVQSQVOLKPRDL 42  
Db 1 MSLNRLILNKAAALRKAHTSMVNFYRGKPVQSQVOLKGRDL 42

RESULT 9  
US-09-791-537-104643  
; Sequence 104643, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
; TITLE OF INVENTION: METHODS OF USE THEREOF  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 104643  
; LENGTH: 354  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-791-537-104643

Query Match 75.6%; Score 164; DB 22; Length 354;  
Best Local Similarity 83.3%; Pred. No. 5.6e-16;  
Matches 35; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 MSLNRLILNKAAALRKAHTSMVNFYRGKPVQSQVOLKPRDL 42  
Db 1 MSLNRLILNKAAALRKAHTSMVNFYRGKPVQSQVOLKGRDL 42

RESULT 10  
PCT-US04-20454-90  
; Sequence 90, Application PC/TUS0420454  
; GENERAL INFORMATION:  
; APPLICANT: Gencia Corporation  
; APPLICANT: Khan, Shaharvar  
; TITLE OF INVENTION: Modified Vectors for Organelle Transfection  
; FILE REFERENCE: 120701-2020  
; CURRENT APPLICATION NUMBER: PCT/US04/20454  
; CURRENT FILING DATE: 2004-07-02  
; PRIOR APPLICATION NUMBER: 60/482,603  
; PRIOR FILING DATE: 2003-06-25  
; NUMBER OF SEQ ID NOS: 585  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 90  
; LENGTH: 366  
; TYPE: PRT  
; ORGANISM: Mus musculus  
PCT-US04-20454-90

Query Match 75.6%; Score 164; DB 1; Length 366;  
Best Local Similarity 83.3%; Pred. No. 5.8e-16;  
Matches 35; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 MSLNRLILNKAAALRKAHTSMVNFYRGKPVQSQVOLKPRDL 42





APPLICANT: Bionomix, Inc.  
APPLICANT: Debe, Derek  
APPLICANT: Danzer, Joseph  
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
FILE REFERENCE: 261/210  
CURRENT APPLICATION NUMBER: US/09/791,537  
CURRENT FILING DATE: 2001-02-22  
NUMBER OF SEQ ID NOS: 153055  
SOFTWARE: Patent in version 3.0  
SEQ ID NO 88687  
LENGTH: 354  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-791-537-88687

Query Match 65.0%; Score 141; DB 28; Length 354;  
Best Local Similarity 69.0%; Pred. No. 2.1e-12;  
Matches 29; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 1 MLSNLRILLNKAALRKAHTSMVNFYRGKPVQSQVQLKPRDL 42  
DB 1 MLFNLRILLNNAAFRNGHFMVNFRCGQPLQNKVQLKGRDL 42

RESULT 17  
US-10-219-051B-12769  
Sequence 12769, Application US/10219051B  
GENERAL INFORMATION:  
APPLICANT: The General Hospital Corporation doing business as Massachusetts General  
APPLICANT: Hospital / Bayer AG  
TITLE OF INVENTION: Nucleotide sequences involved in pain  
FILE REFERENCE: Lea 35693 Foreign Countries  
CURRENT APPLICATION NUMBER: US/10/219,051B  
CURRENT FILING DATE: 2003-05-09  
PRIOR FILING DATE: 2001-08-14  
PRIOR APPLICATION NUMBER: US 60/312,147  
PRIOR FILING DATE: 2001-08-14  
PRIOR APPLICATION NUMBER: US 60/346,382  
PRIOR FILING DATE: 2001-11-01  
PRIOR APPLICATION NUMBER: US 60/333,347  
PRIOR FILING DATE: 2001-11-26  
NUMBER OF SEQ ID NOS: 14715  
SOFTWARE: Perl script  
SEQ ID NO 12769  
LENGTH: 354  
TYPE: PRT  
ORGANISM: Homo sapiens  
PUBLICATION INFORMATION:  
DATABASE ACCESSION NUMBER: Refseq / NP\_000522  
DATABASE ENTRY DATE: 2000-10-31  
US-10-219-051B-12769

Query Match 65.0%; Score 141; DB 28; Length 354;  
Best Local Similarity 69.0%; Pred. No. 2.1e-12;  
Matches 29; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 1 MLSNLRILLNKAALRKAHTSMVNFYRGKPVQSQVQLKPRDL 42  
DB 1 MLFNLRILLNNAAFRNGHFMVNFRCGQPLQNKVQLKGRDL 42

RESULT 18  
US-10-219-051B-14342  
Sequence 14342, Application US/10219051B  
GENERAL INFORMATION:  
APPLICANT: The General Hospital Corporation doing business as Massachusetts General  
APPLICANT: Hospital / Bayer AG  
TITLE OF INVENTION: Nucleotide sequences involved in pain  
FILE REFERENCE: Lea 35693 Foreign Countries  
CURRENT APPLICATION NUMBER: US/10/219,051B  
CURRENT FILING DATE: 2003-05-09  
PRIOR FILING DATE: 2001-08-14  
PRIOR APPLICATION NUMBER: US 60/312,147  
PRIOR FILING DATE: 2001-08-14

PRIOR APPLICATION NUMBER: US 60/346,382  
PRIOR FILING DATE: 2001-11-01  
PRIOR APPLICATION NUMBER: US 60/333,347  
PRIOR FILING DATE: 2001-11-26  
NUMBER OF SEQ ID NOS: 14715  
SOFTWARE: Perl script  
SEQ ID NO 14342  
LENGTH: 354  
TYPE: PRT  
ORGANISM: Homo sapiens  
PUBLICATION INFORMATION:  
DATABASE ACCESSION NUMBER: SWISS-Prot / P00480  
DATABASE ENTRY DATE: 2002-06-15  
US-10-219-051B-14342

Query Match 65.0%; Score 141; DB 28; Length 354;  
Best Local Similarity 69.0%; Pred. No. 2.1e-12;  
Matches 29; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 1 MLSNLRILLNKAALRKAHTSMVNFYRGKPVQSQVQLKPRDL 42  
DB 1 MLFNLRILLNNAAFRNGHFMVNFRCGQPLQNKVQLKGRDL 42

RESULT 19  
US-09-791-537-53980  
Sequence 53980, Application US/09791537  
GENERAL INFORMATION:  
APPLICANT: Bionomix, Inc.  
APPLICANT: Debe, Derek  
APPLICANT: Danzer, Joseph  
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
FILE REFERENCE: 261/210  
CURRENT APPLICATION NUMBER: US/09/791,537  
CURRENT FILING DATE: 2001-02-22  
NUMBER OF SEQ ID NOS: 153055  
SOFTWARE: Patent in version 3.0  
SEQ ID NO 53980  
LENGTH: 354  
TYPE: PRT  
ORGANISM: Bos taurus  
US-09-791-537-53980

Query Match 59.9%; Score 130; DB 22; Length 354;  
Best Local Similarity 64.3%; Pred. No. 1.1e-10;  
Matches 27; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLSNLRILLNKAALRKAHTSMVNFYRGKPVQSQVQLKPRDL 42  
DB 1 MLFHLRILLNNAALRNGHFMVNFRCGQPLQDKVQLKGRDL 42

RESULT 20  
US-09-791-537-53995  
Sequence 53995, Application US/09791537  
GENERAL INFORMATION:  
APPLICANT: Bionomix, Inc.  
APPLICANT: Debe, Derek  
APPLICANT: Danzer, Joseph  
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
FILE REFERENCE: 261/210  
CURRENT APPLICATION NUMBER: US/09/791,537  
CURRENT FILING DATE: 2001-02-22  
NUMBER OF SEQ ID NOS: 153055  
SOFTWARE: Patent in version 3.0  
SEQ ID NO 53995  
LENGTH: 354  
TYPE: PRT  
ORGANISM: Trachemys scripta elegans  
US-09-791-537-53995

```
Query Match      48.8%; Score 106; DB 22; Length 354;
Best Local Similarity 52.4%; Pred. No. 6.1e-07;
Matches 22; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

QY 1 MLSNRLILLKAAALRKAHTSMVRNFRYKPVQSQVOLKPRDL 42
   ||||| ||||| : : : : : : : : : : : : : : : : : :
Db 1 MLFNRLNLLNAATLRSSKQLVQFRSGQPTQTNILKGRDL 42

RESULT 21
PCT-US02-02814-48
; Sequence 48, Application PC/TUS020202814
; GENERAL INFORMATION:
; APPLICANT: University of Rochester
; APPLICANT: Zauderer, Maurice (U.S. Only)
; APPLICANT: Smith, Ernest S. (U.S. Only)
; TITLE OF INVENTION: Methods of Identifying Regulator Molecules
; FILE REFERENCE: 1821.0080003
; CURRENT APPLICATION NUMBER: PCT/US02/02814
; CURRENT FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: 60/271,423
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/265,880
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/265,589
; PRIOR FILING DATE: 2001-02-02
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Heterologous signal sequence for the mitochondrial matrix
; NAME/KEY: MISC FEATURE
; LOCATION: (7)..(8)
; OTHER INFORMATION: May be any amino acid
; NAME/KEY: MISC FEATURE
; LOCATION: (32)..(32)
; OTHER INFORMATION: May be any amino acid
PCT-US02-02814-48

Query Match      42.4%; Score 92; DB 1; Length 32;
Best Local Similarity 61.3%; Pred. No. 4.1e-06;
Matches 19; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLSNRLILLKAAALRKAHTSMVRNFRYKPV 31
   ||||| ||||| : : : : : : : : : : : :
Db 1 MLFNRLXXLNNAAFRGHNFVNRNFCGQPL 31

RESULT 22
PCT-US02-21677-54
; Sequence 54, Application PC/TUS0221677
; GENERAL INFORMATION:
; APPLICANT: University of Rochester
; APPLICANT: Zauderer, Maurice
; APPLICANT: Wei, Chungwen
; APPLICANT: Smith, Ernest
; TITLE OF INVENTION: Methods of Producing or Identifying Intrabodies in Eukaryotic Cells
; FILE REFERENCE: 1821.009PC05
; CURRENT APPLICATION NUMBER: PCT/US02/21677
; CURRENT FILING DATE: 2003-02-27
; PRIOR APPLICATION NUMBER: 60/298,095
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/271,422
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/263,200
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 60/263,225
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 154
; SOFTWARE: PatentIn version 3.0
```

```
; SEQ ID NO 54
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: signal sequence
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (7)..(8)
; OTHER INFORMATION: Xaa may represent any amino acid
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (32)..(32)
; OTHER INFORMATION: Xaa may represent any amino acid
PCT-US02-21677-54

Query Match      42.4%; Score 92; DB 1; Length 32;
Best Local Similarity 61.3%; Pred. No. 4.1e-06;
Matches 19; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLSNRLILLKAAALRKAHTSMVRNFRYKPV 31
   ||||| ||||| : : : : : : : : : : : :
Db 1 MLFNRLXXLNNAAFRGHNFVNRNFCGQPL 31

RESULT 23
PCT-US04-11481-23
; Sequence 23, Application PC/TUS0411481
; GENERAL INFORMATION:
; APPLICANT: Medimmune, Inc.
; TITLE OF INVENTION: EphA2 and Hypoproliferative Cell Disorders and Epithelial and Endo
; FILE REFERENCE: 10271-058-228
; CURRENT APPLICATION NUMBER: PCT/US04/11481
; CURRENT FILING DATE: 2004-04-20
; PRIOR APPLICATION NUMBER: 60/462,009
; PRIOR FILING DATE: 2003-04-11
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 23
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (7)..(8)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (32)..(32)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
PCT-US04-11481-23

Query Match      42.4%; Score 92; DB 1; Length 32;
Best Local Similarity 61.3%; Pred. No. 4.1e-06;
Matches 19; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLSNRLILLKAAALRKAHTSMVRNFRYKPV 31
   ||||| ||||| : : : : : : : : : : : :
Db 1 MLFNRLXXLNNAAFRGHNFVNRNFCGQPL 31

RESULT 24
PCT-US04-11482-23
; Sequence 23, Application PC/TUS0411482
; GENERAL INFORMATION:
; APPLICANT: Medimmune, Inc.
; TITLE OF INVENTION: EphA2 and Hyperproliferative Cell Disorders
; FILE REFERENCE: 10271-060-228
; CURRENT APPLICATION NUMBER: PCT/US04/11482
; CURRENT FILING DATE: 2004-04-20
; PRIOR APPLICATION NUMBER: 60/462,024
; PRIOR FILING DATE: 2003-04-11
```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: localization signal used to direct intrabody to mitochondrial mat:
; NAME/KEY: VARIANT
; LOCATION: 7, 8, 32
; OTHER INFORMATION: Xaa = Any Amino Acid
PCT-US04-23192-26

Query Match          42.4%; Score 92; DB 1; Length 32;
Best Local Similarity 61.3%; Pred. No. 4.1e-06;
Matches 19; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLSNRLTLNKAALRKAHTSMVRNFRYKPV 31
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 MLFNLRXXLNNAAPRHGHNFVRNFRCGQPL 31

RESULT 27
US-07-916-939-29
; Sequence 29, Application US/07916939
; GENERAL INFORMATION:
; APPLICANT: MARASCO, WAYNE
; APPLICANT: HASELTINE, WILLIAM
; TITLE OF INVENTION: METHOD OF INTRACELLULAR BINDING OF
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
; ADDRESSEE: CUSHMAN
; STREET: 130 WATER STREET
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/916,939
; FILING DATE: 19920717
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: EISENSTEIN, RONALD I.
; REGISTRATION NUMBER: 30628
; REFERENCE/DOCKET NUMBER: 41956
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 523-3400
; TELEFAX: (617) 523-6440
; TELEX: 200291 STRE UR
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
;
US-07-916-939-29

Query Match          42.4%; Score 92; DB 3; Length 32;
Best Local Similarity 61.3%; Pred. No. 4.1e-06;
Matches 19; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLSNRLTLNKAALRKAHTSMVRNFRYKPV 31
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 MLFNLRXXLNNAAPRHGHNFVRNFRCGQPL 31

RESULT 28

```

;; APPLICANT: MARASCO, wayne

FEEDING SYSTEM: FC-

APPLICANT: MARASCO, Wayne

APPLICANT: MHASHILKAR, Abner  
TITLE OF INVENTION: INTRABODY-MEDIATED CONTROL OF IMMUNE REACTIONS  
FILE REFERENCE: 47577 C

CURRENT APPLICATION NUMBER: US/09/522,727D

CURRENT FILING DATE: 2000-03-10

PRIOR APPLICATION NUMBER: PCT/US98/19563

PRIOR FILING DATE: 1998-09-18

PRIOR APPLICATION NUMBER: 60/059,339

PRIOR FILING DATE: 1997-09-19

NUMBER OF SEQ ID NOS: 56

SOFTWARE: PatentIn version 3.1

SEQ ID NO 29

LENGTH: 32

TYPE: PRT

ORGANISM: human

FEATURE:

NAME/KEY: UNSURE

LOCATION: (7) (8) (32)

OTHER INFORMATION: UNSURE

US-09-522-727D-29

Query Match 42.4%; Score 92; DB 19; Length 32;

Best Local Similarity 61.3%; Pred. No. 4.1e-06;

Matches 19; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLSNRLILNKALRKAKHTSMVRNFRYKPV 31

DB 1 MLFNLRXXLLNNAAFRHHGFMVRNFRCGQPL 31

RESULT 32

US-10-052-942-54

Sequence 54, Application US/10052942

GENERAL INFORMATION:

APPLICANT: Zauderer, Maurice

APPLICANT: Smith, Ernest

APPLICANT: Wei, Chungwen

TITLE OF INVENTION: Methods of Producing or Identifying Intrabodies in Eukaryotic Cell

FILE REFERENCE: 1821.0090004

CURRENT APPLICATION NUMBER: US/10/052,942

CURRENT FILING DATE: 2002-01-23

PRIOR APPLICATION NUMBER: 60/298,095

PRIOR FILING DATE: 2001-06-15

PRIOR APPLICATION NUMBER: 60/271,422

PRIOR FILING DATE: 2001-02-27

PRIOR APPLICATION NUMBER: 60/263,200

PRIOR FILING DATE: 2001-01-24

PRIOR APPLICATION NUMBER: 60/263,225

PRIOR FILING DATE: 2001-01-23

NUMBER OF SEQ ID NOS: 154

SOFTWARE: PatentIn version 3.0

SEQ ID NO 54

LENGTH: 32

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: signal sequence

NAME/KEY: UNSURE

LOCATION: (7) (8)

OTHER INFORMATION: Xaa may represent any amino acid

NAME/KEY: UNSURE

LOCATION: (32) (32)

OTHER INFORMATION: Xaa may represent any amino acid

US-10-052-942-54

Query Match 42.4%; Score 92; DB 26; Length 32;

Best Local Similarity 61.3%; Pred. No. 4.1e-06;

Matches 19; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLSNRLILNKALRKAKHTSMVRNFRYKPV 31

DB 1 MLFNLRXXLLNNAAFRHHGFMVRNFRCGQPL 31

RESULT 33

US-10-061-395-48

Sequence 48, Application US/10061395

GENERAL INFORMATION:

APPLICANT: Zauderer, Maurice

APPLICANT: Smith, Ernest S.

TITLE OF INVENTION: Methods of Identifying Regulator Molecules

FILE REFERENCE: 1821.0080003

CURRENT APPLICATION NUMBER: US/10/061,395

CURRENT FILING DATE: 2002-02-04

PRIOR APPLICATION NUMBER: 60/271,423

PRIOR FILING DATE: 2001-02-27

PRIOR APPLICATION NUMBER: 60/265,880

PRIOR FILING DATE: 2001-02-05

PRIOR APPLICATION NUMBER: 60/265,589

PRIOR FILING DATE: 2001-02-02

NUMBER OF SEQ ID NOS: 116

SOFTWARE: PatentIn version 3.1

SEQ ID NO 48

LENGTH: 32

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Heterologous signal sequence for the mitochondrial matrix

NAME/KEY: MISC FEATURE

LOCATION: (7) (8)

OTHER INFORMATION: May be any amino acid

NAME/KEY: MISC FEATURE

LOCATION: (32) (32)

OTHER INFORMATION: May be any amino acid

US-10-061-395-48

Query Match 42.4%; Score 92; DB 26; Length 32;

Best Local Similarity 61.3%; Pred. No. 4.1e-06;

Matches 19; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLSNRLILNKALRKAKHTSMVRNFRYKPV 31

DB 1 MLFNLRXXLLNNAAFRHHGFMVRNFRCGQPL 31

RESULT 34

US-10-823-254-23

Sequence 23, Application US/10823254

GENERAL INFORMATION:

APPLICANT: Kiener, Peter

APPLICANT: Kinch, Michael

APPLICANT: Langermann, Solomon

APPLICANT: Reed, Jennifer

TITLE OF INVENTION: EphA2 and Hyperproliferative Cell Disorders

FILE REFERENCE: 10271-060-999

CURRENT APPLICATION NUMBER: US/10/823,254

CURRENT FILING DATE: 2004-04-12

PRIOR APPLICATION NUMBER: 60/462,024

PRIOR FILING DATE: 2003-04-11

NUMBER OF SEQ ID NOS: 45

SOFTWARE: PatentIn version 3.2

SEQ ID NO 23

LENGTH: 32

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

LOCATION: (7) (8)

OTHER INFORMATION: Xaa can be any naturally occurring amino acid

FEATURE:

NAME/KEY: misc feature

LOCATION: (32) (32)

OTHER INFORMATION: Xaa can be any naturally occurring amino acid

US-10-823-254-23

Query Match 42.4%; Score 92; DB 34; Length 32;

Best Local Similarity 61.3%; Pred. No. 4.1e-06;  
Matches 19; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 MSLNRLILNKAAALRKAHTSMVRNFRYKPV 31  
DB 1 MLFNLRXXLNNAAFRHGHNFVRNFRCGQPL 31

RESULT 35  
US-10-823-259-23  
; Sequence 23, Application US/10823259  
; GENERAL INFORMATION:  
; APPLICANT: Kiener, Peter  
; APPLICANT: Kincer, Michael  
; APPLICANT: Langermann, Solomon  
; TITLE OF INVENTION: EphA2 and Hyperproliferative Cell Disorders and Epithelial and  
; FILE REFERENCE: 10271-058-999  
; CURRENT APPLICATION NUMBER: US/10/823,259  
; CURRENT FILING DATE: 2004-04-12  
; PRIOR APPLICATION NUMBER: 60/462,009  
; PRIOR FILING DATE: 2003-04-11  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 23  
; LENGTH: 32  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (7)..(8)  
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (32)..(32)  
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
US-10-823-259-23

Query Match 42.4%; Score 92; DB 34; Length 32;  
Best Local Similarity 61.3%; Pred. No. 4.1e-06;  
Matches 19; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 MSLNRLILNKAAALRKAHTSMVRNFRYKPV 31  
DB 1 MLFNLRXXLNNAAFRHGHNFVRNFRCGQPL 31

RESULT 36  
US-09-791-537-53985  
; Sequence 53985, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 53985  
; LENGTH: 356  
; TYPE: PRT  
; ORGANISM: Sceloporus undulatus  
US-09-791-537-53985

Query Match 32.3%; Score 70; DB 22; Length 356;  
Best Local Similarity 40.9%; Pred. No. 0.25;  
Matches 18; Conservative 6; Mismatches 18; Indels 2; Gaps 1;

QY 1 MSLNRLILNKAAALRKAHTSMVRNFRY--GKPVQSOVLKPRDL 42  
DB 1 MLFNLRXXLNNAAFRHGHNFVRNFRCGQPL 31

Db 1 MLFNFRSLFSTRNVNKKSHLVQRIRYRHGPPSETPVQLKGRDL 44

RESULT 37  
US-09-708-427-25599  
; Sequence 25599, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 25599  
; LENGTH: 1235  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 1..1235  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc feature  
; LOCATION: 1..1235  
; OTHER INFORMATION: Ceres Seq. ID 1814485  
US-09-708-427-25599

Query Match 28.1%; Score 61; DB 21; Length 1235;  
Best Local Similarity 38.2%; Pred. No. 32;  
Matches 13; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

QY 3 SNLRILNKAAALRKAHTSMVRNFRYKPVQSOVQ 36  
DB 56 SNLRRLSSSTTKRDESLVRNLLVSPQLDIQ 89

RESULT 38  
US-09-708-427-25598  
; Sequence 25598, Application US/09708427  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES  
; FILE REFERENCE: 2750-1243P  
; CURRENT APPLICATION NUMBER: US/09/708,427  
; CURRENT FILING DATE: 2000-11-09  
; NUMBER OF SEQ ID NOS: 85364  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 25598  
; LENGTH: 1247  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 1..1247  
; OTHER INFORMATION: Xaa is any amino acid  
; NAME/KEY: misc feature  
; LOCATION: 1..1247  
; OTHER INFORMATION: Ceres Seq. ID 1814484  
US-09-708-427-25598

Query Match 28.1%; Score 61; DB 21; Length 1247;  
Best Local Similarity 38.2%; Pred. No. 32;  
Matches 13; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

QY 3 SNLRILNKAAALRKAHTSMVRNFRYKPVQSOVQ 36  
DB 68 SNLRRLSSSTTKRDESLVRNLLVSPQLDIQ 101

RESULT 39  
PCT-US02-18153-3

Search completed: December 18, 2004, 03:03:05  
Job time : 340.905 secs

**This Page Blank (uspto)**



Result No.	Score	Query Match	Length	DB	ID	Description
1	164	75.6	366	1	PCT-US04-35137-101	Sequence 101, App
2	164	75.6	366	6	US-10-972-963-101	Sequence 101, App
3	141	65.0	354	1	PCT-US04-35137-100	Sequence 100, App
4	141	65.0	354	6	US-10-972-963-100	Sequence 100, App
5	141	65.0	354	6	US-10-990-328-9578	Sequence 9578, Ap
6	141	65.0	354	6	US-10-990-328-9579	Sequence 9579, Ap
7	53	24.4	435	6	US-10-732-923-2547	Sequence 2547, Ap
8	53	24.4	593	1	PCT-US02-09107B-78319	Sequence 78319, A
9	52	24.0	1903	1	PCT-US04-02460-3	Sequence 3 Appli
10	51	23.5	598	1	PCT-US02-09107B-73330	Sequence 73330, A
11	50.5	23.3	2379	6	US-10-211-028A-7	Sequence 7, Appli
12	49.5	22.8	200	1	PCT-US02-09107B-57678	Sequence 57678, A
13	49.5	22.8	864	6	US-10-732-923-3312	Sequence 3312, Ap
14	49	22.6	353	6	US-10-732-923-15291	Sequence 15291, A
15	49	22.6	446	1	PCT-US02-09107B-55126	Sequence 55126, A
16	48.5	22.4	135	8	US-60-622-712-1944	Sequence 1944, Ap
17	48.5	22.4	295	1	PCT-US02-09107B-42529	Sequence 42529, A
18	48.5	22.4	5032	6	US-10-408-765-26	Sequence 26, Appl
19	48.5	22.4	5037	5	US-09-424-783-4	Sequence 4, Appli
20	48	22.1	274	6	US-10-732-923-18888	Sequence 18888, A
21	48	22.1	417	6	US-10-965-898-67	Sequence 67, Appl
22	48	22.1	513	1	PCT-US02-09107B-61451	Sequence 61451, A
23	48	22.1	731	1	PCT-US02-09107B-49817	Sequence 49817, A
24	48	22.1	770	6	US-10-972-024-235	Sequence 235, App
25	47	21.7	98	6	US-10-220-366A-20931	Sequence 20931, A

```

Db      1 MLFNLRILLNNAAFRNGHNFVNRFCGQPLQNKVQLKGRDL 42


RESULT 5
US-10-990-328-9578
; Sequence 9578, Application US/10990328
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001495
; CURRENT APPLICATION NUMBER: US/10/990,328
; CURRENT FILING DATE: 2004-11-17
; NUMBER OF SEQ ID NOS: 558824
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9578
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-990-328-9578

Query Match          65.0%; Score 141; DB 6; Length 354;
Best Local Similarity 69.0%; Pred. No. 1e-12;
Matches 29; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY      1 MLSNRLILNKAAALKAHTSMVRNFRYKGPKVQSQVOLKPRDL 42
        ||||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1 MLFNLRILLNNAAFRNGHNFVNRFCGQPLQNKVQLKGRDL 42


RESULT 6
US-10-990-328-9579
; Sequence 9579, Application US/10990328
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001495
; CURRENT APPLICATION NUMBER: US/10/990,328
; CURRENT FILING DATE: 2004-11-17
; NUMBER OF SEQ ID NOS: 558824
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9579
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-990-328-9579

Query Match          65.0%; Score 141; DB 6; Length 354;
Best Local Similarity 69.0%; Pred. No. 1e-12;
Matches 29; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY      1 MLSNRLILNKAAALKAHTSMVRNFRYKGPKVQSQVOLKPRDL 42
        ||||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1 MLFNLRILLNNAAFRNGHNFVNRFCGQPLQNKVQLKGRDL 42


RESULT 7
US-10-732-923-2547
; Sequence 2547, Application US/10732923
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; LENGTH: 435

```



```
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 2379
; TYPE: PRT
; ORGANISM: Streptomyces roseosporus
US-10-211-028A-7

Query Match      23.3%; Score 50.5; DB 6; Length 2379;
Best Local Similarity 38.7%; Pred. No. 1.1e+02;
Matches 12; Conservative 8; Mismatches 10; Indels 1; Gaps 1;

QY 12 AALRKAHTSMVRNFRYKGPVQSVQVQLKPRDL 42
      ||| : : : ||| : : : ||| : : : ||| : : :
Db 55 AALLRRHANLRAAFRYER-LQRPVQIIPREV 84

RESULT 12
PCT-US02-09107B-57678
; Sequence 57678, Application PC/TUS0209107B
; GENERAL INFORMATION:
; APPLICANT: Elittra Pharmaceuticals Inc.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034VPC
; CURRENT APPLICATION NUMBER: PCT/US02/09107B
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/342,923
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: 10/072,851
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/362,699
; PRIOR FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 57678
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Enterococcus faecium
PCT-US02-09107B-57678

Query Match      22.8%; Score 49.5; DB 1; Length 200;
Best Local Similarity 40.7%; Pred. No. 10;
Matches 11; Conservative 4; Mismatches 3; Indels 9; Gaps 1;

QY 10 NKAALRKAHTSMVRNFRYKGPVQSVQV 36
      ||| : : : ||| : : : ||| : : : ||| : : :
Db 51 NKAALRKAHTSMVRNFRYKGPVQSVQV 68

RESULT 13
US-10-732-923-3312
; Sequence 3312, Application US/10732923
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 3312
; LENGTH: 864
; TYPE: PRT
; ORGANISM: Sulfolobus solfataricus
US-10-732-923-3312

Query Match      22.8%; Score 49.5; DB 6; Length 864;
Best Local Similarity 28.0%; Pred. No. 52;
Matches 11; Conservative 4; Mismatches 3; Indels 9; Gaps 1;

QY 10 NKAALRKAHTSMVRNFRYKGPVQSVQV 36
      ||| : : : ||| : : : ||| : : : ||| : : :
Db 51 NKAALRKAHTSMVRNFRYKGPVQSVQV 68

RESULT 14
US-10-732-923-15291
; Sequence 15291, Application US/10732923
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 15291
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Bean golden yellow mosaic virus
US-10-732-923-15291

Query Match      22.6%; Score 49; DB 6; Length 353;
Best Local Similarity 50.0%; Pred. No. 23;
Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 24 NFRYKGPVQSVQVQLKPRDL 43
      ||| : : : ||| : : : ||| : : : ||| : : :
Db 287 NCKYKGPVQIKGGIPVILC 306

RESULT 15
PCT-US02-09107B-55126
; Sequence 55126, Application PC/TUS0209107B
; GENERAL INFORMATION:
; APPLICANT: Elittra Pharmaceuticals Inc.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034VPC
; CURRENT APPLICATION NUMBER: PCT/US02/09107B
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/342,923
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: 10/072,851
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/362,699
; PRIOR FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55126
; LENGTH: 446
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
PCT-US02-09107B-55126

Query Match      22.6%; Score 49; DB 1; Length 446;
Best Local Similarity 29.5%; Pred. No. 29;
Matches 13; Conservative 9; Mismatches 14; Indels 8; Gaps 1;

QY 5 LRLLNKAALRKAHTSMVRNFRYKGPVQSVQVQLKPR 40
      ||| : : : ||| : : : ||| : : : ||| : : :
Db 226 LQPLVQKALERVQEAFLAKSLKSGHKTVESYKPVETQALQVQ 269

RESULT 16
US-60-622-712-1944
; Sequence 1944, Application US/60622712
```

```

; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3025
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 5032
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765-26

Query Match      22.4%; Score 48.5; DB 6; Length 5032;
Best Local Similarity 40.0%; Pred. No. 5.1e+02;
Matches 14; Conservative 3; Mismatches 11; Indels 7; Gaps 1;

QY      11 KAALRKAHTS-----MVNFRYKGPVQSQVQLK 38
          ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      3088 KAGLRGFFSASDIEKMWENLRGKVSQARTQVK 3122

RESULT 19
US-09-424-783-4
; Sequence 4, Application US/09424783
; GENERAL INFORMATION:
; APPLICANT: Hakamata, Yasuhiro
; APPLICANT: Nishimura, Seiichihiro
; APPLICANT: Barsoumian, Edward Leon
; TITLE OF INVENTION: Human Type 3 Ryanodine Receptor Protein
; TITLE OF INVENTION: and DNA Molecules Coding Therefor
; FILE REFERENCE: 0652.2000000
; CURRENT APPLICATION NUMBER: US/09/424,783
; CURRENT FILING DATE: 1999-12-01
; PRIOR APPLICATION NUMBER: PCT/EP98/02926
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: DE 197 22 317.6
; PRIOR FILING DATE: 1997-05-28
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 5037
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-09-424-783-4

Query Match      22.4%; Score 48.5; DB 5; Length 5037;
Best Local Similarity 40.0%; Pred. No. 5.1e+02;
Matches 14; Conservative 3; Mismatches 11; Indels 7; Gaps 1;

QY      11 KAALRKAHTS-----MVNFRYKGPVQSQVQLK 38
          ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      3089 KAGLRGFFSASDIEKMWENLRGKVSQARTQVK 3123

RESULT 20
US-10-732-923-18888
; Sequence 18888, Application US/10732923
; GENERAL INFORMATION:
; APPLICANT: Edgeron, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15 (52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 18888
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Nicotiana langsdorffii x Nicotiana sanderae
US-10-732-923-18888

Query Match      22.1%; Score 48; DB 6; Length 274;

```



PCT-US02-09107B-59069  
 ; Sequence 59069, Application PC/TUS0209107B  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Elitra Pharmaceuticals Inc.

```

; PCT-US02-09107B-68798
; Sequence 68798, Application PC/TUS0209107B
; GENERAL INFORMATION:
; APPLICANT: Elitra Pharmaceuticals Inc.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034VPC

```

```
; CURRENT APPLICATION NUMBER: PCT/US02/09107B
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/342,923
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: 10/072,851
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/362,699
; PRIOR FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 68798
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Proteus mirabilis
PCT-US02-09107B-68798

Query Match      21.4%; Score 46.5; DB 1; Length 294;
Best Local Similarity 32.4%; Pred. No. 43;
Matches 11; Conservative 7; Mismatches 9; Indels 7; Gaps 1;

Qy      6 RLLNKAALR-----KAHTSMVRNFRYKGPVQ 32
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      9 RFLFKSAVRGELVTVTETYSILENHHYDPVQ 42

RESULT 29
PCT-US02-09107B-49563
; Sequence 49563, Application PC/TUS0209107B
; GENERAL INFORMATION:
; APPLICANT: Elittra Pharmaceuticals Inc.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034VPC
; CURRENT APPLICATION NUMBER: PCT/US02/09107B
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/342,923
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: 10/072,851
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/362,699
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 49563
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Burkholderia fungorum
PCT-US02-09107B-49563

Query Match      21.4%; Score 46.5; DB 1; Length 367;
Best Local Similarity 33.3%; Pred. No. 55;
Matches 12; Conservative 8; Mismatches 13; Indels 3; Gaps 1;

Qy      6 RLLNKAALRKAIT---SMVRNFRYKGPVQSQVQLK 38
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     110 RVLIVNASSRNDYTCFGEMSKSFLGLKLAQETLERK 145

RESULT 30
PCT-US02-09107B-46555
; Sequence 46555, Application PC/TUS0209107B
; GENERAL INFORMATION:
; APPLICANT: Elittra Pharmaceuticals Inc.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034VPC
; CURRENT APPLICATION NUMBER: PCT/US02/09107B
```

```
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/342,923
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: 10/072,851
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/362,699
; PRIOR FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46555
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Bacillus anthracis
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (14)..(14)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (189)..(189)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (202)..(202)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (254)..(254)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (268)..(268)
; OTHER INFORMATION: X=any amino acid
PCT-US02-09107B-46555

Query Match      21.4%; Score 46.5; DB 1; Length 510;
Best Local Similarity 34.4%; Pred. No. 79;
Matches 11; Conservative 6; Mismatches 6; Indels 9; Gaps 1;

Qy     12 AALRKAHT-----SMVRNFRYKGPVQSQ 34
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     261 ADIRGAHXTGQKVNEKIALVRNVREGEPIINSR 292

RESULT 31
US-10-732-923-12322
; Sequence 12322, Application US/10732923
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 12322
; LENGTH: 1191
; TYPE: PRT
; ORGANISM: Xanthomonas axonopodis pv. citri str. 306
US-10-732-923-12322

Query Match      21.2%; Score 46; DB 6; Length 1191;
Best Local Similarity 33.3%; Pred. No. 2.4e+02;
Matches 15; Conservative 6; Mismatches 14; Indels 10; Gaps 2;

Qy      7 ILLNKA-----ALRKAHTSMVRN-FRYKGPVQSQVQLKPRD 41
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db     638 VLLDTRAMPDVAFVFPALQLEHVSVLRNGARIALPMDRPFQLGPD 682
```



```

Query Match      21.0%; Score 45.5; DB 1; Length 359;
Best Local Similarity 33.3%; Pred.No.75;
Matches 15; Conservative 8; Mismatches 13; Indels 9; Gaps 2;

QY   3  SNLRILLNKA-----ALRKAHTSMVRNFRYGK-----PVQSQVQLK 38
      .:.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|
Db    269 NSMRIRNAADVSLAPLPPHSSSTRNLIPVKVAECLDVGDDGVDK 313

```

```

Query Match      21.0%; Score 45.5; DB 1; Length 462;
Best Local Similarity 31.9%; Pred. No. 99;
Matches 15; Conservative 13; Mismatches 10; Indels 9; Gaps 4;

Qy 2 LSNLRILLNK--AALRKAH-TSMVRNF--RYGK---PVQSQVLKP 39
Db 196 LKNIKVLNLSGDAAGROAHRDALIAHFPEPRIGFCSDCOSLEKNP 242

```

```
RESULT 36
US-10-952-698-43
; Sequence 43, Application US/10952698
; GENERAL INFORMATION:
; APPLICANT: Mack, David
; APPLICANT: Aziz, Natasha
; TITLE OF INVENTION: METHODS OF DIAGNOSIS OF BLADDER CANCER, COMPOSITIONS AND METHODS
; FILE REFERENCE: OF SCREENING FOR MODULATORS OF BLADDER CANCER
; CURRENT APPLICATION NUMBER: US/10/952,698
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 43
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-952-698-43

Query Match      21.0%; Score 45.5; DB 6; Length 513;
Best Local Similarity 36.7%; Pred. No. 1.1e+02;
Matches 11; Conservative 9; Mismatches 9; Indels 1; Gaps 1;

QY 1 MLSNRLINKAALRKAAHTSMVR-NFRYKX 29
Db 66 LLASLLQILWKGGLKQKQHDTLVEYHKYKX 95

RESULT 37
PCT-US02-09107B-71500
; Sequence 71500, Application PC/TUS0209107B
; GENERAL INFORMATION:
; APPLICANT: Elittra Pharmaceuticals Inc.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034VPC
; CURRENT APPLICATION NUMBER: PCT/US02/09107B
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/342,923
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: 10/072,851
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/362,699
; PRIOR FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 71500
; LENGTH: 786
; TYPE: PRT
; ORGANISM: Staphylococcus haemolyticus
PCT-US02-09107B-71500

Query Match      21.0%; Score 45.5; DB 1; Length 786;
Best Local Similarity 32.1%; Pred. No. 1.8e+02;
Matches 9; Conservative 9; Mismatches 7; Indels 3; Gaps 1;

QY 14 LRKAHTSMVRNFRYKPKVQSQVQLKPRD 41
Db 602 IKKSNVSLFKTKTRYNKPL---VAMKVKD 626

RESULT 38
US-10-408-765-1696
; Sequence 1696, Application US/10408765
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
```

```
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; CURRENT APPLICATION NUMBER: US/10/408,765
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3025
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1696
; LENGTH: 1893
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765-1696

Query Match      21.0%; Score 45.5; DB 6; Length 1893;
Best Local Similarity 33.3%; Pred. No. 4.7e+02;
Matches 11; Conservative 8; Mismatches 13; Indels 1; Gaps 1;

QY 6 RILLNKAALRKAAHTSMVRNFR-YGKPVQSQVQL 37
Db 1294 KVLALQAVREAHAEALLRRAEARGHGLQEQQL 1326

RESULT 39
US-11-000-463-454
; Sequence 454, Application US/11000463
; GENERAL INFORMATION:
; APPLICANT: Tang, Y Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Chen, Rui-hong
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Zhou, Ping
; APPLICANT: Cao, Yi-Cheng
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 785CIP4CN
; CURRENT APPLICATION NUMBER: US/11/000,463
; CURRENT FILING DATE: 2004-11-29
; PRIOR APPLICATION NUMBER: 10/291,265
; PRIOR FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: PCT/US01/02623
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/922,279
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/617,746
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/631,451
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 09/633,870
; NUMBER OF SEQ ID NOS: 944
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 454
; LENGTH: 3674
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-000-463-454

Query Match      21.0%; Score 45.5; DB 7; Length 3674;
Best Local Similarity 33.3%; Pred. No. 9.8e+02;
Matches 11; Conservative 8; Mismatches 13; Indels 1; Gaps 1;

QY 6 RILLNKAALRKAAHTSMVRNFR-YGKPVQSQVQL 37
Db 3075 KVLALQAVREAHAEALLRRAEARGHGLQEQQL 3107
```

```

RESULT 40
US-10-931-081A-174
; Sequence 174, Application US/10931081A
; GENERAL INFORMATION:
; APPLICANT: Lough, Tony James
; APPLICANT: Hermseier, Dieter H.
; APPLICANT: Varkonyi-Gasic, Erika
; APPLICANT: Sweetman, Justin
; APPLICANT: Havukkala, Ilkka J
; APPLICANT: Belanger, Helene
; APPLICANT: Forster, Richard L.S.
; APPLICANT: Hudson, Keith R.
; TITLE OF INVENTION: Control of Floral Induction
; FILE REFERENCE: 1084U
; CURRENT APPLICATION NUMBER: US/10/931,081A
; CURRENT FILING DATE: 2004-08-30
; NUMBER OF SEQ ID NOS: 872
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 174
; LENGTH: 203
; TYPE: PRT
; ORGANISM: Cucurbita maxima
US-10-931-081A-174

```

```

Query Match      20.7%; Score 45; DB 6; Length 203;
Best Local Similarity 75.0%; Pred. No. 47;
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      12 AALRKAHTSMVR 23
Db      30 AALTKAHTSLR 41

```

Search completed: December 18, 2004, 03:03:45  
Job time : 16.869 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 18, 2004, 02:21:06 ; Search time 109.821 Seconds  
(without alignments)  
133.926 Million cell updates/sec

Title: US-08-765-244-1

Perfect score: 208

Sequence: 1 MSLNLRILLNKAALRKAHTS.....VNFYKGVQSQLKRDLC 41

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 23Sep04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	197	94.7	43	AAR90584	Aar90584 Rat ornit
2	179	86.1	354	ADD47079	Add47079 Rat Prote
3	179	86.1	354	ADD48634	Add48634 Rat Prote
4	159	76.4	32	ABG30857	Abg30857 Rat ornit
5	125	60.1	354	ADD48636	Add48636 Human Pro
6	125	60.1	354	ADD47081	Add47081 Human Pro
7	108	51.9	32	ABG30856	Abg30856 Human orn
8	108	51.9	258	AAI18445	Aai18445 Protein e
9	105	50.5	32	AAAG64224	Aag64224 ORC pepti
10	92	44.2	32	AAAG48260	Aar48260 Mitochond
11	92	44.2	32	AAW96358	Aaw96358 Mitochond
12	92	44.2	32	ABG92993	Abg92993 Localisat
13	92	44.2	32	ABP56588	Abp56588 Mitochond
14	79.5	38.2	31	ABP22835	Aab22835 Mitochond
15	79.5	38.2	31	AAAB15704	Aab15704 Mitochond
16	58	27.9	165	AAU04933	Aay04933 Mycobacte
17	55.5	26.7	149	AAU23468	Aau23468 Novel hum
18	55.5	26.7	761	AAAB46718	Aab46718 S. cerevi
19	55.5	26.7	1468	ABRS2610	Abrs2610 Protein s
20	55.5	26.7	1468	ADK64154	Adk64154 Disease t
21	54.5	26.2	138	AAAY75466	Aay75466 Neisseria
22	54.5	26.2	138	AAAY75465	Aay75465 Neisseria
23	54	26.0	65	ADH87578	Adh87578 Enterococ
24	52	25.0	457	ABO84066	AbO84066 Pseudomon
25	51.5	24.8	200	ABU29754	Abu29754 Protein e

## ALIGNMENTS

## RESULT 1

AAR90584  
ID AAR90584 standard; protein; 43 AA.

XX AAR90584;

XX DT 25-MAR-2003 (revised)

XX DT 31-OCT-1996 (first entry)

XX DE Rat ornithine transcarbamylase signal peptide.

XX KW promoter; peptide-nucleic acid; cyclised; gene therapy; target;  
XX site-directed mutagenesis; introduction; protein transport.

XX OS Synthetic.

XX PN DE19520815-A1.

XX PD 21-DEC-1995.

XX PF 11-JUN-1995; 95DE-01020815.

XX PR 16-JUN-1994; 94DE-04421079.

XX PA (SEIB/) SEIBEL P.

XX PI Seibel P, Seibel A;

XX DR WPI; 1996-041226/05.

XX PT Replicable and transcriptionally active plasmid carrying signal peptide  
XX for specific target - useful for site directed mutagenesis and molecular  
XX therapy of genetic diseases.

XX PS Disclosure; Col 11; 24pp; German.

XX CC Two modified oligonucleotides (introducing PstI and XhoI sites) were used  
XX to amplify a region of the human mitochondrial (mt) genome contg. the  
XX light strand promoter, mtDNA ori of the heavy strand, CSB's ("conserved  
XX sequence blocks") and a regulation site for DNA replication. Behind this  
XX fragment (5' direction) a synthetic multiple cloning site was introduced,  
XX generating a product with overhanging ends. The synthetic region also  
XX introduced a bidirectional mt transcription termination sequence. The  
XX amplification product, synthetic fragment and pBluescript were ligated  
XX and recombinant plasmid 1 (AAT12315) was produced. Human mt 16S rRNA  
XX (differing from the native RNA only in having a modified nucleotide) was  
XX isolated by PCR from chloramphenicol resistant HeLa cells and inserted  
XX into plasmid 1 to form plasmid 2 (AAT12316). The cloned insert was

26	51.5	24.8	208	7	ADC95772	Adc95772 E. faeciu
27	51.5	24.8	280	3	AAG04153	Aag04153 Arabidops
28	51.5	24.8	283	3	AAG04152	Aag04152 Arabidops
29	51.5	24.8	303	3	AAG26195	Aag26195 Arabidops
30	51.5	24.8	303	3	AAG37586	Aag37586 Arabidops
31	51.5	24.8	306	3	AAG26194	Aag26194 Arabidops
32	51.5	24.8	306	3	AAG37585	Aag37585 Arabidops
33	51	24.5	135	1	AAP50637	Aap50637 Salmon pl
34	50.5	24.3	296	4	ABB64170	Abb64170 Drosophil
35	50.5	24.3	533	6	ABM67661	Abm67661 Phototrab
36	50.5	24.3	755	4	ABBS9259	Abbs9259 Drosophil
37	50	24.0	74	7	ADH88784	Adh88784 Enterococ
38	50	24.0	136	4	AAB74673	Aab74673 Human pro
39	50	24.0	142	7	ADM06026	Adm06026 Human pro
40	50	24.0	193	5	ABB89739	Abb89739 Human pol
41	50	24.0	250	5	ABB99949	Abb99949 Dipeptidy
42	50	24.0	310	4	AAB47188	Aab47188 Human DPP
43	50	24.0	310	5	ABB08994	Abb08994 Human dip
44	50	24.0	310	7	ADD27858	Add27858 Human dip
45	50	24.0	465	4	AAB47189	Aab47189 Human DPP

CC isolated as a Bsal fragment and cyclised using hairpin loop  
 CC oligonucleotides, one of which carried the required signal peptide (the  
 CC present sequence). The cyclised product was purified by treatment with  
 CC exonuclease III. In a modification, the signal peptide was attached after  
 CC cyclisation. The new plasmids were able to impart chloramphenicol  
 CC resistance to otherwise sensitive B lymphocytes and fibroblasts. Similar  
 CC plasmids without a signal peptide could not do this. The plasmids can be  
 CC introduced into eukaryotic cells, esp. for site-directed mutagenesis or  
 CC molecular therapy of genetic diseases, targeting nucleic acid in cells  
 CC or their organelles via the protein transport route. (Updated on 25-MAR-  
 CC 2003 to correct PR field.)  
 XX  
 SQ Sequence 43 AA;

Query Match 94.7%; Score 197; DB 2; Length 43;  
 Best Local Similarity 95.3%; Pred. No. 2.2e-23;  
 Matches 41; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

QY 1 MLSNRLILNKAAALRKAHTSMVRNFRYKPVQS--QLKPRDLC 41  
 Db 1 MLSNRLILNKAAALRKAHTSMVRNFRYKPVQSQVQLKPRDLC 43

RESULT 2  
 ADD47079  
 ID ADD47079 standard; protein; 354 AA.

XX ADD47079;

XX 29-JAN-2004 (first entry)

XX Rat Protein AAA41767, SEQ ID NO 12767.

XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
 XX Chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX Rattus norvegicus.

XX WO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

XX 01-NOV-2001; 2001US-0346382P.

XX 26-NOV-2001; 2001US-0333347P.

XX (GEO) GEN HOSPITAL CORP.

XX (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

XX GENBANK; AAA41767.

XX New composition comprising two or more isolated polypeptides, useful for  
 XX preparing a medicament for treating pain in an animal.

XX Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a

CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 354 AA;

Query Match 86.1%; Score 179; DB 7; Length 354;  
 Best Local Similarity 92.9%; Pred. No. 2.1e-19;  
 Matches 39; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

QY 1 MLSNRLILNKAAALRKAHTSMVRNFRYKPVQS--QLKPRDL 40  
 Db 1 MLSNRLILNKAAALRKAHTSMVRNFRYKPVQSQVQLKGRDL 42

RESULT 3  
 ADD48634

XX ID ADD48634 standard; protein; 354 AA.

XX ADD48634;

XX 29-JAN-2004 (first entry)

XX Rat Protein OWRT, SEQ ID NO 14340.

XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
 XX Chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX Rattus norvegicus.

XX WO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

XX 01-NOV-2001; 2001US-0346382P.

XX 26-NOV-2001; 2001US-0333347P.

XX (GEO) GEN HOSPITAL CORP.

XX (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

XX New composition comprising two or more isolated polypeptides, useful for  
 XX preparing a medicament for treating pain in an animal.

XX Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC that is differentially expressed in neuronal tissue of a first animal

CC invention  
XX  
SQ Sequence 32 AA;  
  
Query Match 76.4%; Score 159; DB 5; Length 32;  
Best Local Similarity 100.0%; Pred. No. 1.7e-17; Indels 0; Gaps 0;  
Matches 32; Conservative 0; Mismatches 0;  
  
QY 1 MLSNLRILLNKAALRKAHTSMVRNFRYGRPVQ 32  
|||||  
DB 1 MLSNLRILLNKAALRKAHTSMVRNFRYGRPVQ 32  
  
RESULT 5  
ADD48636  
ID ID ADD48636 standard; protein; 354 AA.  
XX  
AC ADD48636;  
XX  
DT 29-JAN-2004 (first entry)  
XX  
DE Human Protein P00480, SEQ ID NO 14342.  
XX  
KW Human; pain; neuronal tissue; gene therapy;  
KW spinal segmental nerve injury; chronic constriction injury; CCI;  
KW spared nerve injury; SNI; Chung.  
XX  
OS Homo sapiens.  
XX  
FN WO2003016475-A2.  
XX  
PD 27-FEB-2003.  
XX  
PF 14-AUG-2002; 2002WO-US025765.  
XX  
PR 14-AUG-2001; 2001US-0312147P.  
PR 01-NOV-2001; 2001US-0346382P.  
PR 26-NOV-2001; 2001US-0333347P.  
XX  
PA (GEO ) GEN HOSPITAL CORP.  
PA (FARB ) BAYER AG.  
XX  
PI Woolf C, D'urso D, Befort K, Costigan M;  
XX  
WI WPI; 2003-268312/26.  
XX  
DR GENBANK; P00480.  
XX  
PT New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.  
XX  
PS Claim 1; Page; 1017pp; English.  
XX  
CC The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction

The invention relates to a recombinant vector for expressing a fused protein containing a fused gene in which a base sequence defining a transfer signal peptide to small cellular organs having an extranuclear gene is combined with a base sequence defining the amino acid sequence of a restriction enzyme recognising a defined base sequence. The vector is used for decomposing an extranuclear gene of a nonhuman organism. The present sequence represents the signal peptide of rat mitochondrial ornithine transcarbamylase (MOT) which may be used in the vector of the

The invention discloses a composition comprising two or more isolated rat, or human polynucleotides or a polynucleotide which represents a fragment, derivative or allelic variation of the nucleic acid sequence. Also claimed are a vector comprising the novel polynucleotide, a host cell comprising the vector, a method for identifying a nucleotide sequence which is differentially regulated in an animal subjected to pain and a kit to perform the method, an array, a method for identifying an agent that increases or decreases the expression of the polynucleotide sequence that is differentially expressed in neuronal tissue of a first animal subjected to pain, a method for identifying a compound which regulates the expression of a polynucleotide sequence which is differentially expressed in an animal subjected to pain, a method for identifying a compound that regulates the activity of one or more of the polynucleotides, a method for producing a pharmaceutical composition, a method for identifying a compound or small molecule that regulates the activity in an animal of one or more of the polypeptides given in the



[illegible]

RESULT 8  
AAB18445  
ID AAB18445 standard; protein; 258 AA.

AC AAB18445;

DT 15-JAN-2001 (first entry)

DE Protein encoded by plasmid pUOATP2.

KW	Mitochondrial genome; mitochondrial encephalomyopathy; lactic acidosis;
KW	Leber hereditary optic neuropathy; myoclonic epilepsy;
KW	neurogenic muscular weakness; ataxia; retinitis pigmentosa;
KW	Kearns-Sayre syndrome; Leigh syndrome; Pearson Marrow pancreas syndrome;
KW	aminoglycoside-associated deafness; diabetes; deafness; leukodystrophy;
KW	hypotonia; autism; sudden infant death syndrome; hypoglycemia; leukaemia;
KW	thrombocytopenia; migraine; hearing loss; stroke;
KW	refractory infantile reflux; carnitine deficiency; multiple sclerosis;
KW	blindness; optic atrophy; renal tubular acidosis; cardiomyopathy;
KW	chronic pancreatitis; Atpase 6.

OS Synthetic.

PN WO200053773-A2.

PD 14-SEP-2000.

PF 08-MAR-2000; 2000WO-US006037.

PR 08-MAR-1999; 99US-0123336P.

PA (USSH ) US DEPT HEALTH &amp; HUMAN SERVICES.

PI Zullo SJ, Eisenstadt JM;

DR WPI; 2000-565602/52.

XX XX

in a mitochondrial genome of a cell, useful for treating diabetes with deafness, comprises introducing mitochondrial DNA into the nuclear genome.

Example 1: Fig 3: 49pp: English.

The specification describes a method for functionally complementing one or more defects, mutations, or deletions in a mitochondrial genome of a cell having a nuclear genome. The method is used for treating a disease or disorder that arises from deletion of the protein-encoding genes of the mitochondrial genome. Alternatively, the method is used for treating a disease or disorder that arises from one or more defects, deletions or mutations in mitochondrial genes encoding ribosomes or tRNA for transcription and translation in the mitochondrial. The diseases or disorders that can be ameliorated are mitochondrial encephalomyopathy with

CC neuropathy, myoclonic epilepsy with ragged-red fibres, neurogenic  
CC muscular weakness, ataxia, retinitis pigmentosa, Kearns-Sayre syndrome;  
CC Leigh syndrome, Pearson Marrow pancreas syndrome, aminoglycoside-  
CC associated deafness, diabetes with deafness, leukodystrophy with  
CC hypotonia, autism with seizures, sudden infant death syndrome with  
CC hypoglycemia, leukaemia with maternally inherited thrombocytopenia,  
CC migraines (associated with hearing loss, strokes, or diabetes), early  
CC hearing loss, refractory infantile reflux with carnitine deficiency,  
CC multiple sclerosis with seizures, blindness with optic atrophy and  
CC dystonia, renal tubular acidosis with elevated lactic acid and hypotoni  
CC nonvalvular hypertrophic cardiomyopathy before age 50 and chronic  
CC pancreatitis with stroke-like episodes. The present sequence is encoded

CC by plasmid pUOATP2. Plasmid pUOATP2 comprises a mutant oligomycin-  
CC resistant ATPase 6 mitochondrial gene derived from Chinese hamster ovary  
CC (CHO) cells linked to ornithine transcarbamylase DNA sequence. It is used  
CC for targeting the protein to the mitochondria  
XX  
XX Sequence 258 AA;  
SQ

Query Match 51.9%; Score 108; DB 3; Length 258;  
Best Local Similarity 68.8%; Pred. No. 2.7e-08;  
Matches 22; Conservative 2; Mismatches 8. Indels

QV 1 MLNLRILLNKAALRKAHTSMVRNFRYKGKPVQ 32

Db 1 MLENRLILLNNAAFRNHGHNFVMVRNFRCCGQPLQ 32

RESULT 9	
AAG64224	
ID AAG64224 standard; peptide; 32 AA.	
XX	
AC	
XX AAG64224;	
XX	
DT 19-SEP-2001 (first entry)	
XX	
XX	
DE OTC peptide fragment.	
XX	
KW	
OTC Heat shock protein interacting protein; HSP47; OTC.	

OS Unidentified.

PN JP2001145493-A.

PD 29-MAY-2001.

PF 19-NOV-1999;

PR 19-NOV-1999; 99JP-00330631.

PA (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.

DR WPI; 2001-395263/42.

Using the two-hybrid screening method to prepare proteins which interact with the heat shock protein HSP47.

PS Example 7; Page 12; 26pp; Japanese.

CC The present invention relates to a method for preparing a protein which  
CC interacts with the heat shock protein HSP47. The method involves the two-  
CC hybrid screening method using the HSP47 gene and a mammalian cDNA  
CC library. The HSP47 interacting proteins are useful for the diagnosis and  
CC treatment of diseases caused by an increase or decrease in activity of  
CC HSP47. The present sequence was used in an example from the present  
CC invention

Sequence 32 AA;

Query Match 50.5%; Score 105; DB 4; Length 32;  
Best Local Similarity 68.8%; Pred. No. 6.3e-09;  
Matches 22; Conservative 2; Mismatches 8; Indels

1 MLNLRILLNKAAALRKAHTSMVRNFRYGKPVQ 32

db 1 MLENRLILLNNAAPRNGHNFMVNRFCGQPLQ 32

## RESULT 10

AAK48260  
ID AAR48260 standard: peptide: 32 AA.

AA  
AC AAR48260:

25-MAR-2003 (revised)

```

DT 29-JUL-1994 (first entry)
DE Mitochondrial matrix retention signal.
XX
XX Single chain antibody; sFv; heavy chain; light chain; variable domain;
XX hydrophilic linker; antibodies; targeting;
KW subcellular localisation signal; mitochondrial matrix; retention signal.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FT Misc-difference 7 /note= "not defined"
FT Misc-difference 8 /note= "not defined"
FT Misc-difference 32 /note= "not defined"
XX
XX WO9402610-A1.
XX
XX 03-FEB-1994.
XX
XX 16-JUL-1993; 93WO-US006735.
XX
XX 17-JUL-1992; 92US-00916939.
XX 17-MAR-1993; 93US-00045274.
XX
XX (DAND ) DANA FARBER CANCER INST INC.
XX
XX Marasco WA, Haseltine WA;
XX WPI; 1994-048868/06.
XX
XX Intracellular binding of antigens - by using antibody targetting with
XX vector system, for e.g. tumour suppression.
XX
XX Disclosure; Page 103; 155pp; English.
XX
XX New vector systems comprise a sequence adapted for intracellular delivery
XX and expression contg. a promoter operably linked to an antibody gene
XX encoding an antibody which binds to a specific target antigen. The
XX antibody is esp. a single chain antibody in which the heavy and light
XX chain variable regions are joined via a hydrophilic linker peptide.
XX Localisation sequences are pref. included in the constructs. The sequence
XX AAR48260 is a mitochondrial matrix retention signal. (Updated on 25-MAR-
XX 2003 to correct PN field.)
XX
XX Sequence 32 AA;
XX
XX Query Match 44.2%; Score 92; DB 2; Length 32;
XX Best Local Similarity 61.3%; Pred. No. 7.3e-07;
XX Matches 19; Conservative 2; Mismatches 10; Indels 0; Gaps 0;
XX
QY 1 MLSNRLILNKALRKAKHTSMVNFVRYGKPV 31
XX ||||| ||||| ||||| ||||| |||||
DB 1 MLFNLRXXLNNAAFRGHGFMVNFVRCGQPL 31
XX ||||| ||||| ||||| ||||| |||||

RESULT 11
AAW96358
ID AAW96358 standard; peptide; 32 AA.
XX
XX AAW96358;
XX
XX 19-JUL-1999 (first entry)
XX
XX Mitochondrial matrix localisation signal peptide.
XX
XX Antibody; immune response; modulation; MHC; IRM; receptor; intrabody;
XX major histocompatibility complex; graft rejection;
XX immunomodulatory response molecule; regulation; transplantation;
XX retention signal; localisation signal; golgi apparatus; ER;
XX endoplasmic reticulum.
XX

```

---

```

XX OS Synthetic.
XX
XX Key Location/Qualifiers
FT Misc-difference 7 /note= "Any amino acid"
FT Misc-difference 8 /note= "Any amino acid"
FT Misc-difference 32 /note= "Any amino acid"
XX
XX WO9914353-A2.
XX
XX 25-MAR-1999.
XX
XX 18-SRP-1998; 98WO-US019563.
XX
XX 19-SEP-1997; 97US-0059339P.
XX
XX (DAND ) DANA FARBER CANCER INST INC.
XX
XX Marasco W, Mhashikar A;
XX WPI; 1999-229546/19.
XX
XX Altering the regulation of the immune system.
XX
XX Disclosure; Page 28; 56pp; English.
XX
XX Intracellular binding to a desired target by an intracellularly expressed
XX antibody (i.e. an intrabody) can be used to knock out multiple locuses of
XX immunomodulatory receptor molecules (IRMs), so that the expression of
XX multiple major histocompatibility (MHC) molecules is blocked. This
XX selectively targetting of IRMs, their pathways or components, can be used to
XX selectively regulate the immune system by controlling expression of these
XX molecules and preventing an undesired immune response in a cell. Any
XX component of the MHC pathway or the MHC assembly line or antigen
XX presentation can be targeted. Intrabodies can be used to knock out the
XX immune response in a particular tissue or portion of the body to prepare
XX it for cell or tissue transplantation. Alternatively, an organ for
XX transplantation can be perfused with the intrabody ex vivo. The
XX intrabodies can comprise whole antibodies, heavy chains, Fab' fragments,
XX single-chain antibodies and diabodies. The intrabodies also comprise an
XX intracellular localisation signal to facilitate interception of expressed
XX proteins. For example, if the target was a cell surface receptor, the
XX antibody would comprise a leader sequence and an endoplasmic reticulum
XX (ER) or Golgi apparatus retention signal. This peptide is a localisation
XX sequence for the mitochondrial matrix. For other localisation sequences
XX see AAW96345-W96377
XX
XX Sequence 32 AA;
XX
XX Query Match 44.2%; Score 92; DB 2; Length 32;
XX Best Local Similarity 61.3%; Pred. No. 7.3e-07;
XX Matches 19; Conservative 2; Mismatches 10; Indels 0; Gaps 0;
XX
QY 1 MLSNRLILNKALRKAKHTSMVNFVRYGKPV 31
XX ||||| ||||| ||||| ||||| |||||
DB 1 MLFNLRXXLNNAAFRGHGFMVNFVRCGQPL 31
XX ||||| ||||| ||||| ||||| |||||

RESULT 12
ABG92993
ID ABG92993 standard; peptide; 32 AA.
XX
XX ABG92993;
XX
XX 20-NOV-2002 (first entry)
XX
XX Localisation sequence to direct antibodies to the mitochondria.
XX
XX Regulator; transcription; cell death; phenotype; molecular scaffold;
XX gene therapy; cancer; cardiovascular disease; arrhythmia; heart failure;
XX

```

KW ischaemia; obesity; neurodegenerative disease; Alzheimer's disease;  
 KW bone pathology; dermatologic disease; psoriasis; infection; AIDS;  
 KW acquired immunodeficiency syndrome; cosmetic; wound healing;  
 KW antibiotic transport; drug toxicity; drug resistance; immunobiology;  
 KW inflammation; allergic response; human immunodeficiency virus.

XX Unidentified.

XX WO200262822-A2.

XX 15-AUG-2002.

XX 04-FEB-2002; 2002WO-US002814.

XX 02-FEB-2001; 2001US-0265589P.

PR 05-FEB-2001; 2001US-0265589P.

PR 27-FEB-2001; 2001US-0271423P.

XX (UYRP ) UNIV ROCHESTER.

XX Zauderer M, Smith ES;

XX WPI; 2002-643398/69.

XX Identifying regulator polypeptides which influence target transcriptional  
 PT regulatory regions, useful for treating cancer, comprises introducing  
 PT host cells expressing the polypeptide into a library of polynucleotides.

XX Disclosure; Page 37; 224pp; English.

XX The invention discloses a method for identifying polynucleotides encoding  
 CC a regulator polypeptide, whose expression induces activation of a target  
 CC transcriptional regulatory region in a host cell. The method comprises  
 CC providing a population of eukaryotic host cells capable of expressing the  
 CC polypeptide, introducing into the host cell a library of polynucleotides  
 CC encoding the polypeptides, permitting expression of the polypeptides and  
 CC then recovering them from the host cells. The target transcriptional  
 CC regulatory region is operably associated with a polynucleotide encoding a  
 CC gene product, the expression of which results in host cell death or cause  
 CC the host cells to exhibit a pre-determined modified phenotype and where  
 CC the gene product is expressed upon activation of target transcriptional  
 CC regulatory region. Each candidate regulator polypeptide comprises a  
 CC candidate peptide and a molecular scaffold fused to the peptide so that  
 CC the peptide is displayed on the surface of the candidate regulator  
 CC polypeptide. The methods are useful in selecting and/or screening  
 CC regulator molecules, such as polypeptides, which directly or indirectly  
 CC induce or suppress the transcriptional activation of a target  
 CC transcriptional regulatory region in a eukaryotic host cell. These  
 CC regulator molecules may be used (e.g. in gene therapy) for preventing or  
 CC treating cancers (e.g. breast or ovarian cancer), cardiovascular diseases  
 CC (e.g. arrhythmia, heart failure, ischaemia), obesity, neurodegenerative  
 CC diseases (e.g. Alzheimer's disease), bone pathologies, dermatologic  
 CC diseases (e.g. psoriasis), infections (e.g. viral, bacterial), acquired  
 CC immunodeficiency syndrome (AIDS), in cosmetic applications and in wound  
 CC healing. The method is also useful in screening regulator molecules that  
 CC block antibiotic transport mechanisms, in drug toxicities and drug  
 CC resistance applications and in improving the performance of existing or  
 CC developmental drugs. It may also be used in immunobiology, inflammation,  
 CC allergic response and in biotechnology applications. The sequences  
 CC presented in ABG92946-ABG93029 are examples of regulator polypeptides

XX Sequence 32 AA;

Query Match 44.2%; Score 92; DB 5; Length 32;  
 Best Local Similarity 61.3%; Pred. No. 7.3e-07;  
 Matches 19; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLSNRLRLNKAALRKAHTSMVFNFRGKPV 31

Db 1 MFLNRLXLLNNAAFRHGHFNFWNFRGQPL 31

RESULT 13

ABP56588

ID ABP56588 standard; peptide; 32 AA.

XX ABP56588;

XX 24-MAR-2003 (first entry)

DT Mitochondrial matrix targeting peptide SEQ ID NO:54.

XX Identification; intrabody; eukaryotic cell; immunoglobulin; selection;  
 KW cardiovascular; diminished arrhythmia potential; cardiomyocyte; stroke;  
 KW enhanced contractile property; heart failure; arrhythmia; embolic;  
 KW sarcolemmal calcium cycling; artery; arteriole; angina; atherosclerosis;  
 KW LDL metabolism; HDL metabolism; skin biology; keloid formation.

XX Unidentified.

XX Key Location/Qualifiers

FT Misc-difference 7 /note= "any amino acid"

FT Misc-difference 8 /note= "any amino acid"

FT Misc-difference 32 /note= "any amino acid"

XX WO200286096-A2.

XX 31-OCT-2002.

XX 23-JAN-2002; 2002WO-US001677.

XX 23-JAN-2001; 2001US-0263225P.

XX 24-JAN-2001; 2001US-0263200P.

XX 27-FEB-2001; 2001US-0271422P.

XX 15-JUN-2001; 2001US-0298095P.

XX (UYRP ) UNIV ROCHESTER MEDICAL CENT.

XX Zauderer M, Wei C, Smith ES;

XX WPI; 2003-103408/09.

XX Selecting polynucleotides encoding an intracellular immunoglobulin which  
 PT induces a modified phenotype in a eukaryotic host cell, by introducing  
 PT library of polynucleotides encoding immunoglobulin subunit polypeptides.

XX Disclosure; Page 44; 257pp; English.

XX The present invention describes a method for selecting polynucleotides  
 CC (PNS) encoding an intracellular immunoglobulin molecule or its fragment  
 CC whose expression induces a modified phenotype in a eukaryotic host cell  
 CC (I). The method comprises introducing into (i) a first and second library  
 CC of PNS encoding, through operable association with a transcriptional  
 CC control region, first and second intracellular immunoglobulin subunit  
 CC polypeptides, respectively. The method is useful for selecting  
 CC polynucleotides which encode an intracellular immunoglobulin molecule, or  
 CC fragment. The method is useful e.g. for identifying polynucleotides which  
 CC singly or collectively encode intracellular immunoglobulin molecules, or  
 CC which sensitize host cells to killing by an agent. The method may also be  
 CC used in cardiovascular applications; for screening for diminished  
 CC arrhythmia potential in cardiomyocytes and for enhanced contractile  
 CC properties of cardiomyocytes and diminish heart failure potential; for  
 CC identifying intracellular immunoglobulin molecules that will regulate  
 CC intracellular and sarcolemmal calcium cycling in cardiomyocytes to  
 CC prevent arrhythmias or that will diminish embolic phenomena in arteries  
 CC and arterioles leading to strokes and angina; in screening for decreases  
 CC in atherosclerosis-producing mechanisms to find intracellular  
 CC immunoglobulin molecules that regulate LDL and HDL metabolism; in skin  
 CC biology applications; and in regulating or inhibiting keloid formation.  
 CC ABZ22379 to ABZ22449 and ABP56536 to ABP56618 represent sequences used in  
 CC the exemplification of the present invention

XX Sequence 32 AA;



CC N-end rule pathway. This method for regulating protein stability allows  
CC removal of the antibody after it has bound to its target antigen.  
CC Stabilised recombinant proteins may be used in gene therapy for the  
CC treatment of disorders such as Alzheimer's disease  
XX  
SQ Sequence 31 AA;  
Query Match 38.2%; Score 79.5; DB 3; Length 31;  
Best Local Similarity 61.3%; Pred. No. 6.8e-05;  
Matches 19; Conservative 1; Mismatches 10; Indels 1; Gaps 1;  
QY 1 MLSNRLINKAALRKAKTSVMRNFRYQKPV 31  
Db 1 MLFNLR-XLNNAAFRGHFMWRNFRGGPL 30  
RESULT 16  
AAV04933  
ID AAY04933 standard; protein; 165 AA.  
XX  
AC AAY04933;  
XX  
DT 06-JUL-1999 (first entry)  
XX  
DE Mycobacterium species protein sequence 38P.  
XX  
KW Secreted protein; Mycobacterium; primer; PCR; amplification; probe;  
KW hybridisation; detection; vaccine; immunisation; infection.  
XX  
OS Mycobacterium sp.  
XX  
PN WO9909186-A2.  
XX  
PD 25-FEB-1999.  
XX  
PF 14-AUG-1998; 98WO-FR001813.  
XX  
PR 14-AUG-1997; 97FR-00010404.  
XX  
PR 11-SEP-1997; 97FR-00011325.  
XX  
PA (INSP ) INST PASTEUR.  
XX  
PI Gicquel B, Portnoie D, Lim E, Pelicic V, Guigueno A;  
PI Goguet De La Salmoniere Y;  
XX  
DR WPI; 1999-181045/15.  
DR N-PSDB; AAX34186.  
XX  
PT Mycobacterial DNA vectors containing reporter constructs - for  
PT identifying coding or promoter sequences involved in infection-associated  
PT protein expression.  
PS Claim 32; Fig 38B; 309pp; French.  
XX  
CC Sequences AAY04742-Y05000 and AAY07201-Y07204 represent secreted proteins  
CC from various Mycobacterium species microorganisms. The encoding  
CC nucleotide sequences can be used as primers and probes for methods for  
CC detecting and identifying mycobacteria, especially belonging to the M.  
CC tuberculosis complex. The encoded proteins can be used in vaccines for  
CC immunisation against a bacterial or viral infection  
XX  
SQ Sequence 165 AA;  
Query Match 27.9%; Score 58; DB 2; Length 165;  
Best Local Similarity 39.4%; Pred. No. 1.3;  
Matches 13; Conservative 7; Mismatches 13; Indels 0; Gaps 0;  
QY 6 RILLNKAALRKAKTSVMRNFRYQKPVQSQKPR 38  
Db 119 RVILRCATKXQNSRARTLRPLRLRLRPR 151  
RESULT 17

AAU23468  
ID AAU23468 standard; protein; 149 AA.  
XX  
AC AAU23468;  
XX  
DT 18-DEC-2001 (first entry)  
XX  
DE Novel human enzyme polypeptide #554.  
XX  
KW Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;  
KW ligase; hyperproliferative disorder; immunodeficiency disorder;  
KW autoimmune disorder; neurological disorder; metabolic disorder;  
KW inflammatory disorder; cardiovascular disorder; reproductive disorder;  
KW blood-related disorder; infectious disorder; cytostatic; anti arthritic;  
KW nephrotropic; anticoagulant.  
XX  
OS Homo sapiens.  
XX  
PN WO200155301-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001239.  
XX  
PR 31-JAN-2000; 2000US-017906SP.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0215647P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 23-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 06-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 08-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.

PN WO200075335-A2.

XX PD 14-DEC-2000.  
 XX PF 02-JUN-2000; 2000WO-IB000893.  
 XX PR 02-JUN-1999; 99US-0137120P.  
 XX PR (DECO-) DECODE GENETICS EHF.  
 XX PI Hjorleifsdottir S, Hreggvidason GO, Fridjonsson OH, Avarsson A;  
 PI Kristjansson JK;  
 XX WPI; 2001-061727/07.  
 XX Isolated nucleic acid from bacteriophage RM 378 encoding enzymes useful  
 PT in recombinant DNA technology.  
 XX PS Disclosure; Fig 3A-P; 42pp; English.  
 XX CC This invention describes a novel isolated nucleic molecule (I) comprising  
 CC the genome of bacteriophage RM 378. The invention also describes (1) an  
 CC isolated nucleic acid which encodes a polypeptide obtainable from  
 CC bacteriophage RM 378, or its active derivative or fragment; (2) an  
 CC isolated bacteriophage RM 378 (II); (3) a DNA construct (III) comprising  
 CC operatively linked to a regulatory sequence; (4) a host cell comprising  
 CC (III); and (5) an isolated polypeptide (IV) obtainable from (II), its  
 CC active derivative or fragment. Bacteriophage RM 378 is useful for  
 CC producing thermophilic enzymes useful in DNA research and commercial  
 CC settings (e.g. proteases and lipases used in washing powder, hydrolytic  
 CC enzymes used in bleaching). The isolated nucleic acid molecules and  
 CC vectors are useful in the manufacture of encoded polypeptide, as probes  
 CC for isolating homologous sequences (e.g. from other bacteriophage  
 CC species), as well as for detecting the presence of the bacteriophage in a  
 CC culture of host cells. The polypeptides can be used as a molecular weight  
 CC marker on SDS-PAGE gels or on molecular sieve gel filtration columns.  
 CC Because the host organism of the RM378 bacteriophage is a thermophile,  
 CC the enzymes and proteins of the RM378 bacteriophage are significantly  
 CC more thermostable than those of other (e.g. mesophilic) bacteriophages,  
 CC such as the T4 bacteriophage of *Escherichia coli*. The enhanced stability  
 CC of the enzymes and proteins of RM378 bacteriophage allows their use under  
 CC temperature conditions which would be prohibitive for other enzymes, thus  
 CC increasing the range of conditions which can be employed not only in DNA  
 CC research but also in commercial settings  
 XX SQ Sequence 761 AA;  
 Query Match 26.7%; Score 55.5; DB 4; Length 761;  
 Best Local Similarity 29.4%; Pred. No. 21;  
 Matches 15; Conservative 4; Mismatches 21; Indels 11; Gaps 1;  
 OY 2 LSNLRILL-----NKAALRKAHTSMVRNFRYKPVQSQLKPRDLC 41  
 DB 62 MPNLRCLSLSIQTLMPNPKENKQIVSITLSAYRNISLDSIPENIKPDDLC 112  
 RESULT 19  
 ID ABR52610  
 XX ABR52610 standard; protein; 1468 AA.  
 XX AC ABR52610;  
 XX DT 20-JUN-2003 (first entry)  
 XX DE Protein sequence #SEQ ID 85.  
 XX KW Multiprotein complex; eukaryote; drug target; diagnosis.  
 XX OS Saccharomyces cerevisiae.  
 XX PN EP1258494-A1.  
 XX PD 20-NOV-2002.  
 XX

PF 20-DEC-2001; 2001EP-00130253.  
 XX PR 15-MAY-2001; 2001EP-00111774.  
 XX PA (CELL-) CELLZONE AG.  
 XX PI Bauer A, Gavin A, Grandi P, Krause R, Kruse U, Kuester BD;  
 PI Marzioch M, Schultz JD, Superti-Furga GD;  
 XX WPI; 2003-250078/25.  
 DR N-PSDB; ACC60652.  
 XX New isolated protein complexes useful for diagnosing a disease or  
 PT disorder, or as a target for an active agent of a pharmaceutical,  
 PT preferably a drug target in the treatment or prevention of disease or  
 PT disorder.  
 XX PS Disclosure; SEQ ID NO 85; 17pp + Sequence Listing; English.  
 XX CC The invention relates to multiprotein complexes from eukaryotes. Proteins  
 CC of the invention and DNA sequences encoding them are given in records  
 CC ABR52568-ABR53903 and ACC0610-ACC061944 respectively. The complexes are  
 CC obtainable by using a protein as a bait and isolating the set of proteins  
 CC which is attached thereto from cells. Such protein complexes may comprise  
 CC up to 30 distinct proteins. Protein complexes of the invention are useful  
 CC for diagnosing a disease or disorder, or as a target for an active agent  
 CC of a pharmaceutical, preferably a drug target in the treatment or  
 CC prevention of a disease or disorder. Note: The sequence data for this  
 CC patent is not represented in the printed specification, but is based on  
 CC sequence information supplied by the European Patent Office. The complete  
 CC document is available on CD-ROM  
 XX SQ Sequence 1468 AA;  
 Query Match 26.7%; Score 55.5; DB 6; Length 1468;  
 Best Local Similarity 29.4%; Pred. No. 48;  
 Matches 15; Conservative 4; Mismatches 21; Indels 11; Gaps 1;  
 OY 2 LSNLRILL-----NKAALRKAHTSMVRNFRYKPVQSQLKPRDLC 41  
 DB 533 MPNLRCLSLSIQTLMPNPKENKQIVSITLSAYRNISLDSIPENIKPDDLC 583  
 RESULT 20  
 ID ADK64154  
 XX ADK64154 standard; protein; 1468 AA.  
 XX AC ADK64154;  
 XX DT 06-MAY-2004 (first entry)  
 XX DE Disease treating protein complex-derived protein #1298.  
 XX KW protein complex; drug target; diagnosis.  
 XX OS Unidentified.  
 XX PN EP1338608-A2.  
 XX PD 27-AUG-2003.  
 XX PF 20-DEC-2002; 2002EP-00102902.  
 XX PR 20-DEC-2001; 2001EP-00130253.  
 XX PA (CELL-) CELLZONE AG.  
 XX PI Bauer A, Gavin A, Superti-Furga G, Kuester B, Schultz J;  
 PI Marzioch M, Grandi P, Krause R, Kruse U, Merino A, Bauch A;  
 PI Michon A, Leutwein C, Rick J;  
 XX WPI; 2003-638460/61.  
 DR N-PSDB; ADK64155.  
 DR

XX New proteins and protein complexes from eukaryotes, useful as targets in  
PT drug screening, or in diagnosing or screening for the presence of a  
PT disease or disorder, or a predisposition for developing a disease or  
PT disorder in a subject.  
XX  
XX Disclosure; SEQ ID NO 2595; 13pp; English.  
XX  
XX The invention relates to novel protein complexes comprising a first and a  
CC second protein, or its derivative, fragment, homologue or variant. The  
CC proteins are selected from given protein complexes, which are not defined  
CC in the specification. The variants are encoded by nucleic acids that  
CC hybridize to the nucleic acids encoding the proteins under low stringency  
CC conditions. The protein complexes are useful as targets for an active  
CC agent of a pharmaceutical. These protein complexes are particularly  
CC useful as drugs targets for the treatment or preventing of a disease or  
CC disorder. The complexes and methods above are useful in diagnosing or  
CC screening for the presence of a disease or disorder or a predisposition  
CC for developing a disease or disorder in a subject. These are also useful  
CC in screening for a drug for treatment or prevention of a disease or  
CC disorder. The molecule that modulates the amount, activity or protein  
CC components of the complex is useful for the manufacture of a medicament  
CC for the treatment or prevention of a disease or disorder. This sequence  
CC corresponds to a protein of the invention. (Note: the sequence data for  
CC this patent did not form part of the printed specification but was  
CC obtained from the EPO in electronic format).

XX Sequence 1468 AA;

Query Match 26.7%; Score 55.5; DB 7; Length 1468;  
Best Local Similarity 29.4%; Pred. No. 48;  
Matches 15; Conservative 4; Mismatches 21; Indels 11; Gaps 1;

Qy 2 LSNRIILL-----NKAALRKHAHTSMVNFYRGKPVQSLKPRDLC 41  
Db 533 MPNLRCLSLSIQTLMNPKNKQEVITLSAYRNISLSDSPIENIKPDDLC 583

RESULT 21

AAAY75466  
ID AAAY75466 standard; protein; 138 AA.

AC AAAY75466;

DT 21-MAR-2000 (first entry)

DE Neisseria meningitidis ORF 715 protein sequence SEQ ID NO:2406.

XX Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;  
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;  
KW antibacterial; gene therapy.

XX Neisseria meningitidis.

XX WO9957280-A2.

XX 11-NOV-1999.

XX 30-APR-1999; 99WO-US009346.

XX 01-MAY-1998; 98US-0083758P.

XX 31-JUL-1998; 98US-0094869P.

XX 02-SEP-1998; 98US-0098994P.

XX 02-SEP-1998; 98US-0099062P.

XX 09-OCT-1998; 98US-0103749P.

XX 09-OCT-1998; 98US-0103796P.

XX 25-FEB-1999; 99US-0121528P.

XX (CHIR ) CHIRON CORP.

XX (GENO-) INST GENOMIC RES.

XX Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;  
PI Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;

PI Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;  
PI Tettelin H, Venter JC;  
XX WPI: 2000-062150/05.  
DR N-PSDB; AAZ54228.  
XX  
PT Novel Neisserial polypeptides predicted to be useful antigens for  
PT vaccines and diagnostics.  
XX  
XX Claim 2; Page 1154; 1453pp; English.

XX AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAZ74253 to AAZ75941  
CC represent novel Neisseria meningitis and N. gonorrhoeae polynucleotides  
CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent  
CC PCR primers used in the exemplification of the present invention. The  
CC polypeptides, the polynucleotides, antibodies and compositions of the  
CC invention can be used as vaccines, as diagnostic reagents, and as  
CC immunogenic compositions. The polypeptides can be used in the manufacture  
CC of medicaments for treating or preventing infection due to Neisseria  
CC bacteria (e.g. meningitis and septicemia), to detect the presence of  
CC Neisseria bacteria, or to raise antibodies. They may also be used to  
CC screen for agonists or antagonists, which may themselves have use as  
CC antibacterial agents. The polynucleotides of the invention may also be  
CC used in gene therapy protocols

XX Sequence 138 AA;

Query Match 26.2%; Score 54.5; DB 3; Length 138;

Best Local Similarity 42.9%; Pred. No. 3.9;

Matches 15; Conservative 4; Mismatches 15; Indels 1; Gaps 1;

Qy 6 RILLKAAALRKHAHTSMVNFYRGKPVQSLKPRD 39

Db 26 RYLLARLRLSETWHTAVKLNFRYAGRPKWLGLKYRD 60

RESULT 22

AAAY75465

ID AAAY75465 standard; protein; 138 AA.

AC AAAY75465;

DT 21-MAR-2000 (first entry)

DE Neisseria meningitidis ORF 715 protein sequence SEQ ID NO:2404.

XX Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;  
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;  
KW antibacterial; gene therapy.

XX Neisseria meningitidis.

XX WO9957280-A2.

XX 11-NOV-1999.

XX 30-APR-1999; 99WO-US009346.

XX 01-MAY-1998; 98US-0083758P.

XX 31-JUL-1998; 98US-0094869P.

XX 02-SEP-1998; 98US-0098994P.

XX 02-SEP-1998; 98US-0099062P.

XX 09-OCT-1998; 98US-0103749P.

XX 09-OCT-1998; 98US-0103796P.

XX 25-FEB-1999; 99US-0121528P.

XX (CHIR ) CHIRON CORP.

XX (GENO-) INST GENOMIC RES.

XX Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;  
PI Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;  
PI Tettelin H, Venter JC;



XX WPI: 2000-062150/05.  
DR N-PSDB; AA254227.  
XX  
PT Novel Neisserial polypeptides predicted to be useful antigens for  
PT vaccines and diagnostics.  
XX  
PS Claim 2; Page 1154; 1453pp; English.  
XX  
CC AA253015 to AA254536, AA254577 to AA254615, and AA274253 to AA275941  
CC represent novel *Neisseria meningitidis* and *N. gonorrhoeae* polynucleotides  
CC and polypeptides. AA254537 to AA254576 and AA254616 to AA255473 represent  
CC PCR primers used in the exemplification of the present invention. The  
CC polypeptides, the polynucleotides, antibodies and compositions of the  
CC invention can be used as vaccines, as diagnostic reagents, and as  
CC immunogenic compositions. The polypeptides can be used in the manufacture  
CC of medicaments for treating or preventing infection due to *Neisseria*  
CC bacteria (e.g. meningitis and septicemia), to detect the presence of  
CC *Neisseria* bacteria, or to raise antibodies. They may also be used to  
CC screen for agonists or antagonists, which may themselves have use as  
CC antibacterial agents. The polynucleotides of the invention may also be  
CC used in gene therapy protocols  
XX  
XX Sequence 138 AA;  
XX  
Query Match 26.2%; Score 54.5; DB 3; Length 138;  
Best Local Similarity 42.9%; Pred. No. 3.9;  
Matches 15; Conservative 4; Mismatches 15; Indels 1; Gaps 1;  
XX  
QY 6 RILLNKAALRKHAHTSMVRNFRY-GKPVQSQQLKPD 39  
DB 26 RYLLMRLSETMTAVKLNFRYAGRPKWGLKYRD 60  
XX  
RESULT 23  
ADH87578  
ID ADH87578 standard; protein; 65 AA.  
XX  
AC ADH87578;  
XX  
DT 22-APR-2004 (first entry)  
XX  
DE Enterococcus faecalis polypeptide #2058.  
XX  
XX Enterococcus faecalis infection; transcription regulatory element;  
KW antibacterial.  
XX  
OS Enterococcus faecalis.  
XX  
FN US6617156-B1.  
XX  
PD 09-SEP-2003.  
XX  
PF 13-AUG-1998; 98US-00134000.  
XX  
PR 15-AUG-1997; 97US-0055778P.  
XX  
XX (DOUC/) DOUCETTE-STAMM L A.  
PA (BUSH/) BUSH D.  
XX  
XX Doucette-Stamm LA, Bush D;  
XX  
XX WPI: 2003-895394/82.  
DR N-PSDB; ADH84173.  
XX  
XX New nucleic acid comprising a sequence encoding an *Enterococcus faecalis*  
PT polypeptide, useful for preparing a composition for diagnosing or  
PT treating *E. faecalis* infection.  
XX  
XX Disclosure; SEQ ID NO 5463; 193pp; English.  
PS  
XX The invention relates to *Enterococcus faecalis* polynucleotides and  
CC polypeptides. The invention also relates to a recombinant expression  
XX  
vector comprising a polynucleotide operably linked to a transcription  
CC regulatory element, a cell comprising a recombinant vector, a method for  
CC producing an *E. faecalis* polypeptide, an isolated nucleic acid comprising  
CC a sequence not given in the specification, a recombinant vector  
CC comprising the nucleic acid and a cell comprising the recombinant vector.  
CC The polynucleotides can be used to detect the presence of *E. faecalis* in  
CC a sample. The sequences are useful for preparing a composition for  
CC diagnosing or treating *Enterococcus faecalis* infection. This sequence  
CC represents an *E. faecalis* polypeptide of the invention.  
XX  
XX Sequence 65 AA;  
XX  
Query Match 26.0%; Score 54; DB 7; Length 65;  
Best Local Similarity 36.8%; Pred. No. 1.9;  
Matches 14; Conservative 7; Mismatches 9; Indels 8; Gaps 2;  
XX  
QY 4 NRIILNKAALRKHAHTSMVRNFR----YKPVQSQQLKP 37  
DB 3 HMEVILNQLLSKAH----RNFTSLQVYGEPIYSIIP 36  
XX  
RESULT 24  
ABO84066  
ID ABO84066 standard; protein; 457 AA.  
XX  
AC ABO84066;  
XX  
DT 29-JUL-2004 (first entry)  
XX  
DE Pseudomonas aeruginosa polypeptide #16241.  
XX  
KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.  
XX  
OS Pseudomonas aeruginosa.  
XX  
PN US6551795-B1.  
XX  
PD 22-APR-2003.  
XX  
PF 18-FEB-1999; 99US-00252991.  
XX  
PR 18-FEB-1998; 98US-0074788P.  
XX  
PR 27-JUL-1998; 98US-0094190P.  
XX  
XX (GENO-) GENOME THERAPEUTICS CORP.  
XX  
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;  
XX  
XX WPI: 2003-615309/58.  
DR N-PSDB; ABD17637.  
XX  
XX Novel isolated nucleic acid encoding *Pseudomonas aeruginosa* polypeptide,  
PT useful as molecular targets for diagnostics, prophylaxis and treatment of  
PT pathological conditions resulting from bacterial infection.  
XX  
XX Disclosure; SEQ ID NO 32812; 455pp; English.  
XX  
XX The invention relates to *Pseudomonas aeruginosa* polypeptides and the  
CC polynucleotides encoding them. The sequences are useful in diagnosis and  
CC therapy of pathological conditions, as molecular targets for diagnostics,  
CC prophylaxis and treatment of pathological conditions resulting from a  
CC bacterial infection, for evaluating a compound, such as a component of  
CC for the ability to bind a *P. aeruginosa* nucleic acid, as components of  
CC effective antibacterial targets, as targets for antibacterial drugs,  
CC including anti-*P. aeruginosa* drugs, as templates for recombinant  
CC production of *P. aeruginosa*-derived peptides or polypeptides, as target  
CC components for diagnosis and/or treatment of *P. aeruginosa*-caused  
CC infection, and in detection of *P. aeruginosa* sequences or other sequences  
CC of *Pseudomonas* species using biochip technology. Sequences ABO67826-  
CC ABO84396 represent *P. aeruginosa* polypeptides of the invention. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format from USPTO at  
CC seqdata.uspto.gov/sequence.html



CC nucleic acid is useful for recombinant production of *Candida albicans* -  
CC derived peptides or antisense polypeptides. Pharmaceutical compositions  
CC and vaccines containing the nucleic acid are useful for preventing or  
CC treating *Enterococcus faecium* infections. The present sequence represents  
CC one if the disclosed *E. faecium* proteins.  
XX  
SQ Sequence 208 AA;

Query Match 24.8%; Score 51.5; DB 7; Length 208;  
Best Local Similarity 44.4%; Pred. No. 19;  
Matches 12; Conservative 3; Mismatches 3; Indels 9; Gaps 1;  
QY 10 NKAALRKKAHTSWRNFRYKGPVQSOLK 36  
Db 59 NKA-----QNFYKGPFTPELE 76

## RESULT 27

AAG04153  
ID AAG04153 standard; protein; 280 AA.

XX AAG04153;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 72.

XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.

OS Arabidopsis thaliana.

XX  
PN EP1033405-A2.

XX  
PD 06-SEP-2000.

XX  
PF 25-FEB-2000; 2000EP-00301439.

XX  
PR 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.

PR 23-APR-1999; 99US-0130891P.

PR 28-APR-1999; 99US-0131449P.

PR 30-APR-1999; 99US-0132048P.

PR 30-APR-1999; 99US-0132407P.

PR 04-MAY-1999; 99US-0132484P.

PR 05-MAY-1999; 99US-0132485P.

PR 06-MAY-1999; 99US-0132486P.

PR 06-MAY-1999; 99US-0132487P.

PR 07-MAY-1999; 99US-0132863P.

PR 11-MAY-1999; 99US-0134256P.

PR 14-MAY-1999; 99US-0134218P.

PR 14-MAY-1999; 99US-0134219P.

PR 14-MAY-1999; 99US-0134221P.

PR 14-MAY-1999; 99US-0134370P.

PR 18-MAY-1999; 99US-0134768P.

PR 19-MAY-1999; 99US-0134941P.

PR 20-MAY-1999; 99US-0135124P.

PR 21-MAY-1999; 99US-0135353P.

PR 24-MAY-1999; 99US-0135629P.

PR 25-MAY-1999; 99US-0136021P.

PR 27-MAY-1999; 99US-0136392P.

PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 21-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 26-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.

```
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 08-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 11-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150894P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157553P.
PR 08-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161923P.
PR 28-OCT-1999; 99US-0161931P.
PR 29-OCT-1999; 99US-0162142P.

PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 08-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 11-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150894P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157553P.
PR 08-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161923P.
PR 28-OCT-1999; 99US-0161931P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 24.8%; Score 51.5; DB 3; Length 280;
Best Local Similarity 41.4%; Pred. No. 28;
Matches 12; Conservative 5; Mismatches 5; Indels 7; Gaps 1;

QY 10 NKAALRKAHTSMVRNFRYCKPVQSQLKPR 38
Db 258 NKAATQAHTFXLAN-----EGRUKPR 279

RESULT 28
AAG04152
ID AAG04152 standard; protein; 283 AA.
XX
AC AAG04152;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 71.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 21-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
```

PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139452P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142354P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143342P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.

PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 24.8%; Score 51.5; DB 3; Length 283;  
Best Local Similarity 41.4%; Pred. NO. 28;  
Matches 12; Conservative 5; Mismatches 5; Indels 7; Gaps 1;

QY 10 NKAALRKAHTSMVRNFRYCKPVQSQLKPR 38  
||||| :||| :|  
:|||||

Db 261 NKAATQAHFKLAN-----BGRUKPR 282

RESULT 29

AAG26195

ID AAG26195 standard; protein; 303 AA.

XX AC

XX AAG26195;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 30561.

XX KW Protein identification; signal transduction pathway; metabolic pathway;

XX KW hybridisation assay; genetic mapping; gene expression control; promoter;

XX KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

PR 18-JUN-1999; 99US-0139456P.

PR 18-JUN-1999; 99US-0139457P.

PR 18-JUN-1999; 99US-0139458P.

PR 18-JUN-1999; 99US-0139459P.

PR 18-JUN-1999; 99US-0139460P.

PR 18-JUN-1999; 99US-0139461P.

PR 18-JUN-1999; 99US-0139462P.

PR 18-JUN-1999; 99US-0139463P.

PR 18-JUN-1999; 99US-0139750P.

PR 18-JUN-1999; 99US-0139763P.

PR 21-JUN-1999; 99US-0139817P.

PR 22-JUN-1999; 99US-0139899P.

PR 23-JUN-1999; 99US-0140353P.

PR 23-JUN-1999; 99US-0140354P.

PR 24-JUN-1999; 99US-0140695P.

PR 28-JUN-1999; 99US-0140823P.

PR 29-JUN-1999; 99US-0140991P.

PR 30-JUN-1999; 99US-0141287P.

PR 01-JUL-1999; 99US-0141842P.

PR 01-JUL-1999; 99US-0142154P.

PR 02-JUL-1999; 99US-0142055P.

PR 06-JUL-1999; 99US-0142390P.

PR 08-JUL-1999; 99US-0142803P.

PR 09-JUL-1999; 99US-0142920P.

PR 12-JUL-1999; 99US-0142977P.

PR 13-JUL-1999; 99US-0143542P.

PR 14-JUL-1999; 99US-0143624P.

PR 15-JUL-1999; 99US-0144005P.

PR 16-JUL-1999; 99US-0144085P.

PR 16-JUL-1999; 99US-0144086P.

PR 19-JUL-1999; 99US-0144325P.

PR 19-JUL-1999; 99US-0144331P.

PR 19-JUL-1999; 99US-0144332P.

PR 19-JUL-1999; 99US-0144333P.

PR 19-JUL-1999; 99US-0144334P.

PR 19-JUL-1999; 99US-0144335P.

PR 20-JUL-1999; 99US-0144352P.

PR 20-JUL-1999; 99US-0144632P.

PR 20-JUL-1999; 99US-0144884P.

PR 21-JUL-1999; 99US-0144814P.

PR 21-JUL-1999; 99US-0145086P.

PR 21-JUL-1999; 99US-0145088P.

PR 22-JUL-1999; 99US-0145087P.

PR 22-JUL-1999; 99US-0145089P.

PR 22-JUL-1999; 99US-0145192P.

PR 23-JUL-1999; 99US-0145145P.

PR 23-JUL-1999; 99US-0145218P.

PR 23-JUL-1999; 99US-0145224P.

PR 26-JUL-1999; 99US-0145276P.

PR 27-JUL-1999; 99US-0145913P.

PR 27-JUL-1999; 99US-0145918P.

PR 27-JUL-1999; 99US-0145919P.

PR 28-JUL-1999; 99US-0145951P.

PR 02-AUG-1999; 99US-0146386P.

PR 02-AUG-1999; 99US-0146388P.

PR 02-AUG-1999; 99US-0146389P.

PR 03-AUG-1999; 99US-0147038P.

PR 04-AUG-1999; 99US-0147204P.

PR 04-AUG-1999; 99US-0147302P.

PR 05-AUG-1999; 99US-0147192P.

PR 05-AUG-1999; 99US-0147260P.

PR 06-AUG-1999; 99US-0147303P.

PR 06-AUG-1999; 99US-0147416P.

PR 09-AUG-1999; 99US-0147493P.

PR 09-AUG-1999; 99US-0147935P.

PR 10-AUG-1999; 99US-0148171P.

PR 11-AUG-1999; 99US-0148319P.

PR 12-AUG-1999; 99US-0148341P.

PR 13-AUG-1999; 99US-0148565P.

PR 13-AUG-1999; 99US-0148684P.

PR 16-AUG-1999; 99US-0149368P.

PR 17-AUG-1999; 99US-0149175P.

```
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 23-AUG-1999; 99US-0150566P.
PR 23-AUG-1999; 99US-0150884P.
PR 26-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151108P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158322P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 24.8%; Score 51.5; DB 3; Length 303;
Best Local Similarity 41.4%; Pred. No. 30;
Matches 12; Conservative 5; Mismatches 5; Indels 7; Gaps 1;

QY 10 NKAALRKAKHTSMVRNFRYGPVQSOLKPR 38
DB 281 NKAATKQAKHTFKLAN-----EGRUKPR 302

RESULT 30
AAG37586
ID AAG37586 standard; protein; 303 AA.
XX
```

```
AC AAG37586;
XX 18-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 46237.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 18-MAY-1999; 99US-0134370P.
PR 19-MAY-1999; 99US-0134768P.
PR 20-MAY-1999; 99US-0134941P.
PR 21-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
```

PR 18-JUN-1999; 99US-0139463P.  
 PR 18-JUN-1999; 99US-0139750P.  
 PR 18-JUN-1999; 99US-0139763P.  
 PR 21-JUN-1999; 99US-0139817P.  
 PR 22-JUN-1999; 99US-0139899P.  
 PR 23-JUN-1999; 99US-0140353P.  
 PR 23-JUN-1999; 99US-0140354P.  
 PR 24-JUN-1999; 99US-0140695P.  
 PR 28-JUN-1999; 99US-0140823P.  
 PR 29-JUN-1999; 99US-0140991P.  
 PR 30-JUN-1999; 99US-0141287P.  
 PR 01-JUL-1999; 99US-0141842P.  
 PR 01-JUL-1999; 99US-0142154P.  
 PR 02-JUL-1999; 99US-0142055P.  
 PR 08-JUL-1999; 99US-0142390P.  
 PR 08-JUL-1999; 99US-0142803P.  
 PR 08-JUL-1999; 99US-0142820P.  
 PR 12-JUL-1999; 99US-0142977P.  
 PR 13-JUL-1999; 99US-0143542P.  
 PR 14-JUL-1999; 99US-0143624P.  
 PR 15-JUL-1999; 99US-0144005P.  
 PR 16-JUL-1999; 99US-0144085P.  
 PR 16-JUL-1999; 99US-0144086P.  
 PR 19-JUL-1999; 99US-0144325P.  
 PR 19-JUL-1999; 99US-0144331P.  
 PR 19-JUL-1999; 99US-0144332P.  
 PR 19-JUL-1999; 99US-0144333P.  
 PR 19-JUL-1999; 99US-0144334P.  
 PR 19-JUL-1999; 99US-0144335P.  
 PR 20-JUL-1999; 99US-0144352P.  
 PR 20-JUL-1999; 99US-0144632P.  
 PR 20-JUL-1999; 99US-0144884P.  
 PR 21-JUL-1999; 99US-0144814P.  
 PR 21-JUL-1999; 99US-0145086P.  
 PR 21-JUL-1999; 99US-0145088P.  
 PR 22-JUL-1999; 99US-0145085P.  
 PR 22-JUL-1999; 99US-0145087P.  
 PR 22-JUL-1999; 99US-0145089P.  
 PR 22-JUL-1999; 99US-0145192P.  
 PR 23-JUL-1999; 99US-0145145P.  
 PR 23-JUL-1999; 99US-0145218P.  
 PR 23-JUL-1999; 99US-0145224P.  
 PR 26-JUL-1999; 99US-0145276P.  
 PR 27-JUL-1999; 99US-0145913P.  
 PR 27-JUL-1999; 99US-0145918P.  
 PR 27-JUL-1999; 99US-0145919P.  
 PR 28-JUL-1999; 99US-0145951P.  
 PR 02-AUG-1999; 99US-0146386P.  
 PR 02-AUG-1999; 99US-0146388P.  
 PR 02-AUG-1999; 99US-0146389P.  
 PR 03-AUG-1999; 99US-0147038P.  
 PR 04-AUG-1999; 99US-0147204P.  
 PR 05-AUG-1999; 99US-0147302P.  
 PR 05-AUG-1999; 99US-0147182P.  
 PR 06-AUG-1999; 99US-0147260P.  
 PR 06-AUG-1999; 99US-0147303P.  
 PR 06-AUG-1999; 99US-0147416P.  
 PR 09-AUG-1999; 99US-0147493P.  
 PR 09-AUG-1999; 99US-0147935P.  
 PR 10-AUG-1999; 99US-0148171P.  
 PR 11-AUG-1999; 99US-0148319P.  
 PR 12-AUG-1999; 99US-0148341P.  
 PR 13-AUG-1999; 99US-0148565P.  
 PR 13-AUG-1999; 99US-0148684P.  
 PR 16-AUG-1999; 99US-0149368P.  
 PR 17-AUG-1999; 99US-0149175P.  
 PR 18-AUG-1999; 99US-0149426P.  
 PR 20-AUG-1999; 99US-0149722P.  
 PR 20-AUG-1999; 99US-0149723P.  
 PR 20-AUG-1999; 99US-0149929P.  
 PR 23-AUG-1999; 99US-0149902P.  
 PR 23-AUG-1999; 99US-0149910P.  
 PR 25-AUG-1999; 99US-0150566P.

PR 26-AUG-1999; 99US-0150884P.  
 PR 27-AUG-1999; 99US-0151065P.  
 PR 27-AUG-1999; 99US-0151066P.  
 PR 27-AUG-1999; 99US-0151080P.  
 PR 30-AUG-1999; 99US-0151303P.  
 PR 31-AUG-1999; 99US-0151438P.  
 PR 01-SEP-1999; 99US-0151930P.  
 PR 07-SEP-1999; 99US-0152363P.  
 PR 10-SEP-1999; 99US-0153070P.  
 PR 13-SEP-1999; 99US-0153758P.  
 PR 16-SEP-1999; 99US-0154018P.  
 PR 16-SEP-1999; 99US-0154039P.  
 PR 20-SEP-1999; 99US-0154779P.  
 PR 22-SEP-1999; 99US-0155139P.  
 PR 23-SEP-1999; 99US-0155486P.  
 PR 24-SEP-1999; 99US-0155659P.  
 PR 28-SEP-1999; 99US-0156458P.  
 PR 29-SEP-1999; 99US-0156596P.  
 PR 04-OCT-1999; 99US-0157117P.  
 PR 05-OCT-1999; 99US-0157753P.  
 PR 06-OCT-1999; 99US-0157865P.  
 PR 07-OCT-1999; 99US-0158029P.  
 PR 08-OCT-1999; 99US-0158232P.  
 PR 12-OCT-1999; 99US-0158369P.  
 PR 13-OCT-1999; 99US-0159293P.  
 PR 13-OCT-1999; 99US-0159294P.  
 PR 13-OCT-1999; 99US-0159295P.  
 PR 14-OCT-1999; 99US-0159329P.  
 PR 14-OCT-1999; 99US-0159330P.  
 PR 14-OCT-1999; 99US-0159331P.  
 PR 14-OCT-1999; 99US-0159637P.  
 PR 14-OCT-1999; 99US-0159638P.  
 PR 18-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 23-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161992P.  
 PR 28-OCT-1999; 99US-0161993P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 24.8%; Score 51.5; DB 3; Length 303;  
 Best Local Similarity 41.4%; Pred. No. 30;

Matches 12; Conservative 5; Mismatches 5; Indels 7; Gaps 1;

OY 10 NKAALKAHTSMVRNFRYKGPVQSOLKPR 38  
 ||||| :||| :  
 Db 281 NKAATKQHTFKLAN-----EGRKPR 302

RESULT 31  
 AAG26194  
 ID AAG26194 standard; protein; 306 AA.

XX AAG26194;  
 AC AAG26194;  
 XX 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 30560.  
 XX Protein identification; signal transduction pathway; metabolic pathway;  
 KW





PR	07-SEP-1999;	99US-0152363P.	PD	06-SEP-2000.	
PR	10-SEP-1999;	99US-0153070P.	XX		
PR	13-SEP-1999;	99US-0153758P.	PF	25-FEB-2000;	2000EP-00301439.
PR	15-SEP-1999;	99US-0154018P.	XX		
PR	16-SEP-1999;	99US-0154039P.	PR	25-FEB-1999;	99US-0121825P.
PR	20-SEP-1999;	99US-0154779P.	PR	05-MAR-1999;	99US-0123180P.
PR	22-SEP-1999;	99US-0155139P.	PR	09-MAR-1999;	99US-0123548P.
PR	23-SEP-1999;	99US-0155486P.	PR	23-MAR-1999;	99US-0125788P.
PR	24-SEP-1999;	99US-0155659P.	PR	25-MAR-1999;	99US-0126264P.
PR	28-SEP-1999;	99US-0156458P.	PR	29-MAR-1999;	99US-0126785P.
PR	29-SEP-1999;	99US-0156596P.	PR	01-APR-1999;	99US-0127462P.
PR	04-OCT-1999;	99US-0157117P.	PR	06-APR-1999;	99US-0128234P.
PR	05-OCT-1999;	99US-0157753P.	PR	08-APR-1999;	99US-0128714P.
PR	06-OCT-1999;	99US-0157865P.	PR	16-APR-1999;	99US-0129845P.
PR	07-OCT-1999;	99US-0158029P.	PR	19-APR-1999;	99US-0130077P.
PR	08-OCT-1999;	99US-0158232P.	PR	21-APR-1999;	99US-0130449P.
PR	12-OCT-1999;	99US-0158369P.	PR	23-APR-1999;	99US-0130510P.
PR	13-OCT-1999;	99US-0159293P.	PR	28-APR-1999;	99US-0130891P.
PR	13-OCT-1999;	99US-0159294P.	PR	30-APR-1999;	99US-0131449P.
PR	14-OCT-1999;	99US-0159329P.	PR	30-APR-1999;	99US-0132048P.
PR	14-OCT-1999;	99US-0159330P.	PR	04-MAY-1999;	99US-0132407P.
PR	14-OCT-1999;	99US-0159331P.	PR	05-MAY-1999;	99US-0132484P.
PR	14-OCT-1999;	99US-0159637P.	PR	06-MAY-1999;	99US-0132485P.
PR	18-OCT-1999;	99US-0159584P.	PR	06-MAY-1999;	99US-0132487P.
PR	21-OCT-1999;	99US-0160741P.	PR	07-MAY-1999;	99US-0132863P.
PR	21-OCT-1999;	99US-0160767P.	PR	11-MAY-1999;	99US-0134218P.
PR	21-OCT-1999;	99US-0160768P.	PR	14-MAY-1999;	99US-0134219P.
PR	21-OCT-1999;	99US-0160770P.	PR	14-MAY-1999;	99US-0134221P.
PR	21-OCT-1999;	99US-0160814P.	PR	14-MAY-1999;	99US-0134370P.
PR	21-OCT-1999;	99US-0160815P.	PR	18-MAY-1999;	99US-0134768P.
PR	22-OCT-1999;	99US-0160980P.	PR	19-MAY-1999;	99US-0134941P.
PR	22-OCT-1999;	99US-0160981P.	PR	20-MAY-1999;	99US-0135124P.
PR	22-OCT-1999;	99US-0160989P.	PR	21-MAY-1999;	99US-0135353P.
PR	25-OCT-1999;	99US-0161405P.	PR	21-MAY-1999;	99US-0135629P.
PR	25-OCT-1999;	99US-0161406P.	PR	25-MAY-1999;	99US-0136021P.
PR	25-OCT-1999;	99US-0161359P.	PR	27-MAY-1999;	99US-0136392P.
PR	26-OCT-1999;	99US-0161360P.	PR	28-MAY-1999;	99US-0136782P.
PR	26-OCT-1999;	99US-0161361P.	PR	01-JUN-1999;	99US-0137222P.
PR	28-OCT-1999;	99US-0161920P.	PR	03-JUN-1999;	99US-0137528P.
PR	28-OCT-1999;	99US-0161992P.	PR	04-JUN-1999;	99US-0137502P.
PR	28-OCT-1999;	99US-0161993P.	PR	07-JUN-1999;	99US-0137724P.
PR	29-OCT-1999;	99US-0162142P.	PR	08-JUN-1999;	99US-0138094P.
PR			PR	10-JUN-1999;	99US-0138540P.
PR			PR	10-JUN-1999;	99US-0138847P.
PR			PR	14-JUN-1999;	99US-0139119P.
PR			PR	16-JUN-1999;	99US-0139452P.
PR			PR	16-JUN-1999;	99US-0139453P.
PR			PR	17-JUN-1999;	99US-0139492P.
Qy	10 NKAALRKAAHTSMVNFYRGKPVOSQLKPR 38		PR	18-JUN-1999;	99US-0139454P.
Db	284 NKAATKAHTFKLAN-----EGRLKPR 305		PR	18-JUN-1999;	99US-0139455P.
			PR	18-JUN-1999;	99US-0139456P.
			PR	18-JUN-1999;	99US-0139457P.
			PR	18-JUN-1999;	99US-0139458P.
			PR	18-JUN-1999;	99US-0139459P.
			PR	18-JUN-1999;	99US-0139460P.
			PR	18-JUN-1999;	99US-0139461P.
			PR	18-JUN-1999;	99US-0139462P.
			PR	18-JUN-1999;	99US-0139463P.
			PR	18-JUN-1999;	99US-0139750P.
			PR	18-JUN-1999;	99US-0139763P.
			PR	21-JUN-1999;	99US-0139817P.
			PR	23-JUN-1999;	99US-0139899P.
			PR	23-JUN-1999;	99US-0140353P.
			PR	23-JUN-1999;	99US-0140354P.
			PR	24-JUN-1999;	99US-0140695P.
			PR	24-JUN-1999;	99US-0140823P.
			PR	28-JUN-1999;	99US-0140921P.
			PR	30-JUN-1999;	99US-0141287P.
			PR	01-JUL-1999;	99US-0141842P.
			PR	01-JUL-1999;	99US-0142154P.
			PR	02-JUL-1999;	99US-0142055P.

Query Match 24.8; Score 51.5; DB 3; Length 306;  
Best Local Similarity 41.4; Pred. No. 31;  
Matches 12; Conservative 5; Mismatches 5; Indels 7; Gaps 1;  
Qy 10 NKAALRKAAHTSMVNFYRGKPVOSQLKPR 38  
Db 284 NKAATKAHTFKLAN-----EGRLKPR 305  
RESULT 32  
AAG37585  
ID AAG37585 standard; protein; 306 AA.  
AC AAG37585;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 46236.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX

PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144684P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 21-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 28-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 09-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 10-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.

PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 24.8%; Score 51.5; DB 3; Length 305;

Best Local Similarity 41.4%; Pred. No. 31;

Matches 12; Conservative 5; Mismatches 5; Indels 7; Gaps 1;

QY 10 NKAALRKHAHTSMVRNFRYKGVQSLKPR 38  
||||:||||:||||:  
DB 284 NKAATKQAHTFKLAN-----EGRCLKPR 305

RESULT 33

AAP50637

ID AAP50637 standard; protein; 135 AA.

XX AAP50637;

XX AC

XX AAP50637;

DT 25-MAR-2003 (revised)

DT 15-JAN-1992 (first entry)

XX Salmon pleiomelanocorticotrophin.

DE DE

XX Salmon hypophysis cerebri hormones; ACTH; gamma-LPH; beta-MSH; CLIP;

KW Beta-endorphin; salmon gonadotrophin; E.coli.

XX Oncothryncus keta.

OS JP60176588-A.

PN JP60176588-A.

XX 10-SEP-1985.

PD 10-SEP-1985.

XX 24-FEB-1984; 84JP-00032700.

XX 24-FEB-1984; 84JP-00032700.

XX (SEK ) SEIKAGAKU KOGYO CO LTD.

XX WPI; 1985-261180/42.  
DR N-PSDB; AAN50500.  
XX  
PT DNA plasmid and its prepn. - having de-ox:ribose nucleotide sequence to  
PT code salmon-hypophysis cerebri-hormone (precursor).  
XX  
PS Disclosure; Page 562; 9pp; Japanese.  
XX  
CC The gene product may be expressed from an E.coli plasmid expression  
CC system for the mass production of various kinds of salmon hypophysis  
CC cerebri hormones eg. ACTH; beta-gamma-LPH; alpha-beta-MSH; CLIP; beta  
CC endorphin etc. (Updated on 25-MAR-2003 to correct PA field.)  
XX  
SQ Sequence 135 AA;  
XX  
Query Match 24.5%; Score 51; DB 1; Length 135;  
Best Local Similarity 36.4%; Pred. No. 14;  
Matches 8; Conservative 7; Mismatches 7; Indels 0; Gaps 0;  
XX  
QY 16 KAHTSMVRNFRYKGPVQSLKP 37  
Db 4 KRHSYSMEHFRWKGPKIGHKRP 25  
XX  
RESULT 34  
ABB64170  
ID ABB64170 standard; protein; 296 AA.  
XX  
AC ABB64170;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster polypeptide SEQ ID NO 19302.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
XX 27-SEP-2001.  
XX  
XX 23-MAR-2001; 2001WO-US009231.  
XX  
XX 23-MAR-2000; 2000US-0191637P.  
PR 11-JUL-2000; 2000US-00614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
PI Venter JC, Adams M, Li PWD, Myers EW;  
XX  
DR WPI; 2001-656860/75.  
DR N-PSDB; ABL06273.  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signaling and cell-cell  
PT interactions..  
XX  
PS Disclosure; SEQ ID NO 19302; 21pp + Sequence Listing; English.  
XX  
CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-  
CC ABB72072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pt\_sequences

SQ Sequence 296 AA;  
XX  
Query Match 24.3%; Score 50.5; DB 4; Length 296;  
Best Local Similarity 44.4%; Pred. No. 42;  
Matches 12; Conservative 5; Mismatches 9; Indels 1; Gaps 1;  
XX  
QY 2 LSNLRILLNKAALRKAH-TSMVFNPRY 27  
Db 76 LEFISILFKKAHLKSHRLKVKNFY 102  
XX  
RESULT 35  
ABM67661  
ID ABM67661 standard; protein; 533 AA.  
XX  
AC ABM67661;  
XX  
DT 20-NOV-2003 (first entry)  
XX  
DE Photorhabdus luminescens protein sequence #758.  
XX  
KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;  
KW detection; food; gene expression; plant; animal; microorganism; toxin;  
KW antibiotic; biopesticide; virulence factor; disease model; plague;  
KW whooping cough.  
XX  
OS Photorhabdus luminescens.  
XX  
PN WO200294867-A2.  
XX  
PD 28-NOV-2002.  
XX  
PF 07-FEB-2002; 2002WO-IB003040.  
XX  
PR 07-FEB-2001; 2001FR-00001659.  
XX  
PA (INSP ) INST PASTEUR.  
PA (CNRS ) CNRS CENT NAT RECH SCI.  
XX  
PI Duchaud E, Taourit S, Glaeser P, Frangeul L, Kunst F, Danchin A;  
PI Buchrieser C;  
XX  
XX WPI; 2003-148459/14.  
XX  
XX Genomic sequence of Photorhabdus luminescens and encoded polypeptides,  
PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.  
XX  
PS Claim 2; SEQ ID NO 758; 1205pp; French.  
XX  
CC The invention relates to the isolation of genes and their encoded  
CC proteins from Photorhabdus luminescens. The isolated sequences are  
CC sources of probes and primers for detecting the genome of P. luminescens  
CC and related species; to study polymorphisms; for gene analysis and for  
CC detection/amplification of the genes. Antibodies (Ab) raised against the  
CC polypeptides encoded by the genes are used for detection/identification  
CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that  
CC carry a gene-containing vector are used to select compounds that  
CC modulate, regulate, induce or inhibit expression of the genes in plants,  
CC animals or microorganisms other than P. luminescens and are able to alter  
CC response or sensitivity to toxins and antibiotics produced by P.  
CC luminescens. Cells transformed to express the genes are useful for  
CC recombinant production of the proteins, particularly toxins and  
CC antibacterials useful as insecticides, bactericides and fungicides. The  
CC genes, proteins, vectors containing the genes and Ab are also useful  
CC therapeutically (to treat microbial infection by bacteria or fungi that  
CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as  
CC biopesticides. Other uses of the genes and the proteins are as virulence  
CC factors and for identifying targets of human diseases for which P.  
CC luminescens is a model (particularly plague and whooping cough). This  
CC sequence represents one of the isolated P. luminescens proteins  
XX  
SQ Sequence 533 AA;

```
Query Match      24.3%; Score 50.5; DB 6; Length 533;
Best Local Similarity 47.8%; Pred. No. 87;
Matches 11; Conservative 6; Mismatches 5; Indels 1; Gaps 1;

QY 2 LSNRLILNKA-ALRKAHTSMVR 23
   ||:::| | | | | | | | | |
Db 285 LSSIRLVSTNGALRKQHVSMIK 307

RESULT 36
ABB59259
ID ABB59259 standard; protein; 755 AA.
XX
AC ABB59259;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 4569.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL03362.
XX
XX
New isolated nucleic acid detection reagent for detecting 1000 or more
genes from Drosophila and for elucidating cell signaling and cell-cell
interactions.
XX
PS Disclosure; SEQ ID NO 4569; 21pp + Sequence Listing; English.
XX
The invention relates to an isolated nucleic acid detection reagent
capable of detecting 1000 or more genes from Drosophila. The invention is
useful in developmental biology and in elucidating cell signalling and
cell-cell interactions in higher eukaryotes for the development of
insecticides, therapeutics and pharmaceutical drugs. The invention
discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
ABB72072). The sequence data for this patent did not form part of the
printed specification, but was obtained in electronic format directly
from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 755 AA;

Query Match      24.3%; Score 50.5; DB 4; Length 755;
Best Local Similarity 42.5%; Pred. No. 1.3e+02;
Matches 17; Conservative 7; Mismatches 11; Indels 5; Gaps 3;

QY 2 LSNRLILNKAALRKAHTSMVRNFRY--GKPVOSQLKPRD 39
   ||:::| | | | | | | | | |
Db 11 LSDLRRELE--VLKKAHFEELDLHFLVGTGQP-EAEAKPRD 47

RESULT 37
ADH88784
ID ADH88784 standard; protein; 74 AA.
XX
AC ADH88784;

Query Match      24.0%; Score 50; DB 7; Length 74;
Best Local Similarity 37.8%; Pred. No. 9.5;
Matches 14; Conservative 6; Mismatches 9; Indels 8; Gaps 2;

QY 5 LRILNKAALRKAHTSMVRNFR-----YKPVOSQLKAP 37
   ||:::| | | | | | | | | |
Db 13 LKLILKQLLLSKAH----RNFTSPQVYGFPGYGIIP 45

RESULT 38
AAB74673
ID AAB74673 standard; protein; 136 AA.
XX
AC AAB74673;
XX
DT 12-JUN-2001 (first entry)
XX
DE Human protease and protease inhibitor PP1M-6.
XX
KW Human; protease; protease inhibitor; protease and protease inhibitor;
KW PPM1; identification; diagnosis; anti-human immunodeficiency virus; HIV;
KW antidiabetic; immunostimulant; immunomodulator; antiinflammatory;
KW antithyroid; immunosuppressive; nephrotropic; antitumor; thyromimetic;
KW cytostatic; antibacterial; fungicide; protozoacide; antiarteriosclerotic;
KW antiatherosclerotic; antipsoriatic; virucide; hepatotropic; gene therapy;
KW autoimmune disorder; inflammatory disorder; AIDS; DiGeorge's syndrome;
KW severe combined immunodeficiency disease; SCID; Chediak-Higashi syndrome;
```



```

XX 19-MAY-2000; 2000US-0205515P.
PR (HUMA-) HUMAN GENOME SCI INC.
XX
XX Birse CE, Rosen CA;
PI
XX
DR WPI: 2002-122018/16.
DR N-PSDB; ABL90148.
XX
XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
PT prevention of neural, immune system, muscular, reproductive,
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
PT disorders.
XX
PS Claim 11; SEQ ID NO 2115; 2081pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (ABL89449-ABL90853) and proteins
CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 193 AA;
Query Match 24.0%; Score 50; DB 5; Length 193;
Best Local Similarity 44.0%; Pred. No. 30;
Matches 11; Conservative 6; Mismatches 6; Indels 2; Gaps 1;
QY 17 AHTSMVRNF--RYGKPVQSLKPRD 39
DB 134 AHTSILLSEFLVRAGRPYDLQIYPQE 158

```

Search completed: December 18, 2004, 02:45:40  
Job time : 114.821 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 18, 2004, 02:37:41 ; Search time 27.3333 Seconds  
(without alignments)  
99.477 Million cell updates/sec

Title: US-08-765-244-1  
Perfect score: 208  
Sequence: 1 MLSNLRILLNKAALRKAHTS.....VNRFRYKPVQSQLKPRDL 41

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/prodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/prodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/prodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/prodata/1/iaa/PCUTUS\_COMB.pep.\*  
6: /cgn2\_6/prodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	125	60.1	354	US-09-538-092-833	Sequence 833, App
2	92	44.2	32	US-08-373-190-29	Sequence 29, Appl
3	92	44.2	32	US-08-438-190A-29	Sequence 29, Appl
4	92	44.2	32	US-08-350-215-29	Sequence 29, Appl
5	92	44.2	32	US-09-287-145A-29	Sequence 29, Appl
6	92	44.2	32	US-09-556-111-29	Sequence 29, Appl
7	60	28.8	205	US-09-270-767-34638	Sequence 34638, A
8	60	28.8	205	US-09-270-767-49855	Sequence 49855, A
9	55.5	26.7	761	US-09-585-858-27	Sequence 27, Appl
10	54	26.0	65	US-09-134-000C-5463	Sequence 5463, Ap
11	53.5	25.7	226	US-09-270-767-56794	Sequence 56794, A
12	53.5	25.7	602	US-09-270-767-41564	Sequence 41564, A
13	52	25.0	457	US-09-252-991A-32812	Sequence 32812, A
14	51.5	24.8	208	US-09-107-532A-5399	Sequence 5399, Ap
15	50	24.0	74	US-09-134-000C-6669	Sequence 6669, Ap
16	50	24.0	310	US-09-794-236-4	Sequence 4, Appli
17	50	24.0	593	US-08-900-927-1	Sequence 1, Appli
18	50	24.0	593	US-09-191-279-1	Sequence 1, Appli
19	50	24.0	593	US-09-334-476-1	Sequence 1, Appli
20	49.5	23.8	475	US-09-543-681A-4698	Sequence 4698, Ap
21	49	23.6	547	US-09-252-991A-25705	Sequence 25705, A
22	49	23.6	550	US-09-489-039A-7879	Sequence 7879, Ap
23	48.5	23.3	274	US-09-252-991A-29653	Sequence 29653, A
24	48	23.1	39	US-08-428-488-13	Sequence 13, Appl
25	48	23.1	89	US-09-732-210-98	Sequence 98, Appl
26	48	23.1	382	US-09-248-796A-19559	Sequence 19559, A
27	48	23.1	567	US-09-188-811-2	Sequence 2, Appli

28	47.5	22.8	96	4	US-09-513-999C-7299	Sequence 7299, Ap
29	47.5	22.8	297	4	US-09-543-681A-6846	Sequence 6846, Ap
30	47	22.6	20	5	PCT-US95-07543-5	Sequence 5, Appli
31	47	22.6	91	4	US-09-732-210-195	Sequence 195, App
32	47	22.6	318	4	US-09-270-767-36225	Sequence 36225, A
33	47	22.6	318	4	US-09-270-767-51442	Sequence 51442, A
34	47	22.6	481	4	US-09-130-491-8	Sequence 8, Appli
35	47	22.6	735	3	US-08-539-205A-2	Sequence 2, Appli
36	47	22.6	735	4	US-09-392-163A-2	Sequence 2, Appli
37	47	22.6	905	3	US-09-369-364A-9	Sequence 9, Appli
38	47	22.6	1050	3	US-09-428-711A-16	Sequence 16, Appli
39	47	22.6	1247	4	US-09-803-286A-2	Sequence 2, Appli
40	47	22.6	2539	3	US-09-413-814-42	Sequence 42, Appl
41	46.5	22.4	1475	4	US-09-538-092-1160	Sequence 1160, Ap
42	46	22.1	78	2	US-08-327-362-3	Sequence 3, Appli
43	46	22.1	78	3	US-09-158-565-3	Sequence 3, Appli
44	46	22.1	193	4	US-09-489-039A-7517	Sequence 7517, Ap
45	46	22.1	331	4	US-09-803-286A-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1  
US-09-538-092-833  
; Sequence 833, Application US/09538092  
; Patent No. 6753314  
; GENERAL INFORMATION:  
; APPLICANT: Giot, Loic  
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same  
; FILE REFERENCE: 15966-542  
; CURRENT APPLICATION NUMBER: US/09/538, 092  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: 60/127,352  
; PRIOR FILING DATE: 1999-04-01  
; PRIOR APPLICATION NUMBER: 60/178,965  
; PRIOR FILING DATE: 2000-02-01  
; NUMBER OF SEQ ID NOS: 1387  
; SOFTWARE: CurapatSeqformat Version 0.9  
; SEQ ID NO 833  
; LENGTH: 354  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (0)...(0)  
; OTHER INFORMATION: Polypeptide Accession Number P00480  
US-09-538-092-833

Query Match 60.1%; Score 125; DB 4; Length 354;  
Best Local Similarity 66.7%; Pred. No. 2.3e-11;  
Matches 28; Conservative 3; Mismatches 9; Indels 2; Gaps 1;

QY 1 MLSNLRILLNKAALRKAHTSMVNRFRYKPVQSQLKPRDL 40  
DB 1 MFLNLRILLNKAALRKAHTSMVNRFRYKPVQSQLKPRDL 42

RESULT 2  
US-08-373-190-29  
; Sequence 29, Application US/08373190  
; Patent No. 5851829  
; GENERAL INFORMATION:  
; APPLICANT: MARASCO, WAYNE  
; APPLICANT: HASELTINE, WILLIAM  
; TITLE OF INVENTION: METHOD OF INTRACELLULAR BINDING OF PROTEINS  
; NUMBER OF SEQUENCES: 79  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN  
; STREET: 130 WATER STREET  
; CITY: BOSTON  
; STATE: MA

```
/ COUNTRY: US
/ ZIP: 02109
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq Version 1.5
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/373,190
/ FILING DATE: 17-JAN-1995
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US93/06735
/ FILING DATE: 16-JUL-1993
/ ATTORNEY/AGENT INFORMATION:
/ NAME: RESNICK, DAVID S
/ REGISTRATION NUMBER: 34,235
/ REFERENCE/DOCKET NUMBER: 41956-PCT-US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617-523-3400
/ TELEFAX: 617-523-6440
/ TELEX: STRE UR 2002
/ INFORMATION FOR SEQ ID NO: 29:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 32 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: N-terminal
/ ORIGINAL SOURCE:
/ US-08-373-190-29

Query Match 44.2%; Score 92; DB 2; Length 32;
Best Local Similarity 61.3%; Pred. No. 2e-07;
Matches 19; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLSNRLINKAALRKAAHTSMVNFYRGKPV 31
Db 1 MLFNLRXXLNNAAFRHGHFMVNFRCGQPL 31

RESULT 3
US-08-438-190A-29
/ Sequence 29, Application US/08438190A
/ Patent No. 5965371
/ GENERAL INFORMATION:
/ APPLICANT: MARASCO, WAYNE
/ APPLICANT: HASELTINE, WILLIAM
/ TITLE OF INVENTION: METHOD OF INTRACELLULAR BINDING OF
/ TITLE OF INVENTION: PROTEINS
/ NUMBER OF SEQUENCES: 78
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
/ ADDRESSEE: CUSHMAN
/ STREET: 130 WATER STREET
/ CITY: BOSTON
/ STATE: MASSACHUSETTS
/ COUNTRY: US
/ ZIP: 02109
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/438,190A
/ FILING DATE:
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: EISENSTEIN, RONALD I.

Query Match 44.2%; Score 92; DB 3; Length 32;
Best Local Similarity 61.3%; Pred. No. 2e-07;
Matches 19; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLSNRLINKAALRKAAHTSMVNFYRGKPV 31
Db 1 MLFNLRXXLNNAAFRHGHFMVNFRCGQPL 31

RESULT 4
US-08-350-215-29
/ Sequence 29, Application US/08350215
/ Patent No. 6004940
/ GENERAL INFORMATION:
/ APPLICANT: MARASCO, WAYNE A.
/ APPLICANT: RICHARDSON, JENNIFER
/ TITLE OF INVENTION: INTRACELLULAR TARGETING OF ENDOGENOUS
/ TITLE OF INVENTION: PROTEINS
/ NUMBER OF SEQUENCES: 58
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
/ ADDRESSEE: CUSHMAN
/ STREET: 130 WATER STREET
/ CITY: BOSTON
/ STATE: MASSACHUSETTS
/ COUNTRY: US
/ ZIP: 02109
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/350,215
/ FILING DATE: 12-DEC-1994
/ CLASSIFICATION: 514
/ ATTORNEY/AGENT INFORMATION:
/ NAME: EISENSTEIN, RONALD I.
/ REGISTRATION NUMBER: 30628
/ REFERENCE/DOCKET NUMBER: 41956-CP3
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (617) 523-3400
/ TELEFAX: (617) 523-6440
/ TELEX: 200291 STRE UR
/ INFORMATION FOR SEQ ID NO: 29:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 32 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ US-08-350-215-29

Query Match 44.2%; Score 92; DB 3; Length 32;
Best Local Similarity 61.3%; Pred. No. 2e-07;
Matches 19; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLSNRLINKAALRKAAHTSMVNFYRGKPV 31
Db 1 MLFNLRXXLNNAAFRHGHFMVNFRCGQPL 31
```

RESULT 5  
US-09-287-145A-29  
; Sequence 29, Application US/09287145A  
; Patent No. 6072036  
; GENERAL INFORMATION:  
; APPLICANT: MARASCO, WAYNE  
; TITLE OF INVENTION: METHOD OF INTRACELLULAR BINDING OF  
; TITLE OF INVENTION: PROTEINS  
; NUMBER OF SEQUENCES: 78  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DAVID G. CONLIN, DIKE, BRONSTEIN, ROBERTS &  
; ADDRESSEE: CUSHMAN  
; STREET: 130 WATER STREET  
; CITY: BOSTON  
; STATE: MASSACHUSETTS  
; COUNTRY: US  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/287,145A  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/438,190  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: EISENSTEIN, RONALD I.  
; REGISTRATION NUMBER: 30628  
; REFERENCE/DOCKET NUMBER: 41956  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 523-3400  
; TELEFAX: (617) 523-6440  
; TELEX: 200291 STRE UR  
; INFORMATION FOR SEQ ID NO: 29:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 32 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 29:  
US-09-287-145A-29  
Query Match 44.2%; Score 92; DB 3; Length 32;  
Best Local Similarity 61.3%; Pred. No. 2e-07;  
Matches 19; Conservative 2; Mismatches 10; Indels 0; Gaps 0;  
QY 1 MLSNRLINKAALRKAHTSMVRNFRYKPV 31  
DB 1 MLFNLEXXLNNAAFRHGHNFMVRNFRGQPL 31  
RESULT 6  
US-09-556-111-29  
; Sequence 29, Application US/09556111  
; Patent No. 6329173  
; GENERAL INFORMATION:  
; APPLICANT: MARASCO, WAYNE  
; TITLE OF INVENTION: METHOD OF INTRACELLULAR BINDING OF  
; TITLE OF INVENTION: PROTEINS  
; NUMBER OF SEQUENCES: 78  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DAVID G. CONLIN, DIKE, BRONSTEIN, ROBERTS &  
; ADDRESSEE: CUSHMAN  
; STREET: 130 WATER STREET  
; CITY: BOSTON  
; STATE: MASSACHUSETTS  
; COUNTRY: US  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/287,145A  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/438,190  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: EISENSTEIN, RONALD I.  
; REGISTRATION NUMBER: 30628  
; REFERENCE/DOCKET NUMBER: 41956  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 523-3400  
; TELEFAX: (617) 523-6440  
; TELEX: 200291 STRE UR  
; INFORMATION FOR SEQ ID NO: 29:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 32 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-09-287-145A-29  
Query Match 44.2%; Score 92; DB 3; Length 32;  
Best Local Similarity 61.3%; Pred. No. 2e-07;  
Matches 19; Conservative 2; Mismatches 10; Indels 0; Gaps 0;  
QY 1 MLSNRLINKAALRKAHTSMVRNFRYKPV 31  
DB 1 MLFNLEXXLNNAAFRHGHNFMVRNFRGQPL 31  
RESULT 7  
US-09-270-767-34638  
; Sequence 34638, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 34638  
; LENGTH: 205  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
; FEATURE:  
; OTHER INFORMATION: Xaa means any amino acid  
US-09-270-767-34638  
Query Match 28.8%; Score 60; DB 4; Length 205;  
Best Local Similarity 42.1%; Pred. No. 0.2;  
Matches 16; Conservative 5; Mismatches 13; Indels 4; Gaps 1;  
QY 8 LLNKAALRKAHTSMVRN----FRYKPVOSQLKPRDLC 41  
DB 66 LINVAHFKAHTILIRNLRLFFSYKSHVKSDDKIXFLC 103  
RESULT 8  
US-09-270-767-49855  
; Sequence 49855, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/556,111  
FILING DATE: 21-Apr-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/438,190  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: EISENSTEIN, RONALD I.  
REGISTRATION NUMBER: 30628  
REFERENCE/DOCKET NUMBER: 41956  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 523-3400  
TELEFAX: (617) 523-6440  
TELEX: 200291 STRE UR  
INFORMATION FOR SEQ ID NO: 29:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 29:  
US-09-556-111-29  
Query Match 44.2%; Score 92; DB 3; Length 32;  
Best Local Similarity 61.3%; Pred. No. 2e-07;  
Matches 19; Conservative 2; Mismatches 10; Indels 0; Gaps 0;  
QY 1 MLSNRLINKAALRKAHTSMVRNFRYKPV 31  
DB 1 MLFNLEXXLNNAAFRHGHNFMVRNFRGQPL 31  
RESULT 7  
US-09-270-767-34638  
; Sequence 34638, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 34638  
; LENGTH: 205  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
; FEATURE:  
; OTHER INFORMATION: Xaa means any amino acid  
US-09-270-767-34638  
Query Match 28.8%; Score 60; DB 4; Length 205;  
Best Local Similarity 42.1%; Pred. No. 0.2;  
Matches 16; Conservative 5; Mismatches 13; Indels 4; Gaps 1;  
QY 8 LLNKAALRKAHTSMVRN----FRYKPVOSQLKPRDLC 41  
DB 66 LINVAHFKAHTILIRNLRLFFSYKSHVKSDDKIXFLC 103  
RESULT 8  
US-09-270-767-49855  
; Sequence 49855, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094



;  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 32812  
; LENGTH: 457  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (451)  
; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.  
US-09-252-991A-32812

Query Match 25.0%; Score 52; DB 4; Length 457;  
Best Local Similarity 34.4%; Pred. No. 9.6;  
Matches 11; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

QY 7 ILINKAALRKHAHTSMVRNRYGKPVQSQLKPR 38  
DB 176 VALDAVRRRLVRFRGLGRPAQLYVETR 207

RESULT 14  
US-09-107-532A-5399  
; Sequence 5399, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 7310  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street  
; CITY: Waltham  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02354  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD/ROM ISO9660  
; COMPUTER: PC  
; OPERATING SYSTEM: <Unknown>  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/107,532A  
; FILING DATE: 30-Jun-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/085,598  
; FILING DATE: 14 May 1998  
; APPLICATION NUMBER: 60/051571  
; FILING DATE: July 2, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ariniello, Pamela Deneke  
; REGISTRATION NUMBER: 40,489  
; REFERENCE/DOCKET NUMBER: GTC-012  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (781)893-5007  
; TELEFAX: (781)893-8277  
; INFORMATION FOR SEQ ID NO: 5399:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 208 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: YES  
; ORIGINAL SOURCE:  
; ORGANISM: Enterococcus faecium

;  
; FEATURE:  
; NAME/KEY: misc-feature  
; LOCATION: (B) LOCATION 1...208  
; SEQUENCE DESCRIPTION: SEQ ID NO: 5399:  
US-09-107-532A-5399

Query Match 24.8%; Score 51.5; DB 4; Length 208;  
Best Local Similarity 44.4%; Pred. No. 4.5;  
Matches 12; Conservative 3; Mismatches 3; Indels 9; Gaps 1;

QY 10 NKAALRKHAHTSMVRNRYGKPVQSQLK 36  
DB 59 NKA-----QNFYRKPTPELE 76

RESULT 15  
US-09-134-000C-6669  
; Sequence 6669, Application US/09134000C  
; Patent No. 6617156  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 032796-032  
; CURRENT APPLICATION NUMBER: US/09/134,000C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/055,778  
; PRIOR FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 6812  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6669  
; LENGTH: 74  
; TYPE: PRT  
; ORGANISM: Enterococcus faecalis  
US-09-134-000C-6669

Query Match 24.0%; Score 50; DB 4; Length 74;  
Best Local Similarity 37.8%; Pred. No. 2.2;  
Matches 14; Conservative 6; Mismatches 9; Indels 8; Gaps 2;

QY 5 LRILNKALRKHAHTSMVRNFR---YGKPVQSQLK 37  
DB 13 LRILNKALLLSKAH---RNFTSPQVYGEYGSIIIP 45

RESULT 16  
US-09-794-236-4  
; Sequence 4, Application US/09794236  
; Patent No. 6337069  
; GENERAL INFORMATION:  
; APPLICANT: Grouzmann, Eric  
; APPLICANT: Lacroix, Jean-Silvain  
; APPLICANT: Monod, Michel  
; TITLE OF INVENTION: Method of Treating Rhinitis and Sinusitis  
; FILE REFERENCE: 81985/276823  
; CURRENT APPLICATION NUMBER: US/09/794,236  
; CURRENT FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 4  
; LENGTH: 310  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-794-236-4

Query Match 24.0%; Score 50; DB 3; Length 310;  
Best Local Similarity 44.0%; Pred. No. 12;  
Matches 11; Conservative 6; Mismatches 6; Indels 2; Gaps 1;

QY 17 AHTSMVRNF--RYGKPVQSQLKPRD 39  
DB 251 AHTSILLSFLVRAGKPYDLQIYPQE 275

RESULT 17  
US-08-900-927-1  
; Sequence 1, Application US/08900927  
; Patent No. 5840537  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Lal, Preeti  
; TITLE OF INVENTION: NEW VESICLE TRANSPORT PROTEIN  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Dr.  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/900,927  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/191,279  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0350 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 593 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: MMLR20T01  
; CLONE: 475485  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0350 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 593 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: MMLR20T01  
; CLONE: 475485  
US-08-900-927-1

Query Match 24.0%; Score 50; DB 2; Length 593;  
Best Local Similarity 31.1%; Pred. No. 27;  
Matches 14; Conservative 9; Mismatches 12; Indels 10; Gaps 2;

QY 4 NLRILNKALRKHTSMVNFY-----GKPVQSQLKPRD 39  
DB 421 NLAKLIQHNV-QAHSSLIRNLEQLGTVNPGSGTSSRLREPE 464

RESULT 18  
US-09-191-279-1  
; Sequence 1, Application US/09191279  
; Patent No. 5981192  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Lal, Preeti  
; TITLE OF INVENTION: NEW VESICLE TRANSPORT PROTEIN  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Dr.  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304

Query Match 24.0%; Score 50; DB 2; Length 593;  
Best Local Similarity 31.1%; Pred. No. 27;  
Matches 14; Conservative 9; Mismatches 12; Indels 10; Gaps 2;

QY 4 NLRILNKALRKHTSMVNFY-----GKPVQSQLKPRD 39  
DB 421 NLAKLIQHNV-QAHSSLIRNLEQLGTVNPGSGTSSRLREPE 464

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/191,279  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/900,927  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0350 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 593 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: MMLR20T01  
CLONE: 475485  
US-09-191-279-1

Query Match 24.0%; Score 50; DB 2; Length 593;  
Best Local Similarity 31.1%; Pred. No. 27;  
Matches 14; Conservative 9; Mismatches 12; Indels 10; Gaps 2;

QY 4 NLRILNKALRKHTSMVNFY-----GKPVQSQLKPRD 39  
DB 421 NLAKLIQHNV-QAHSSLIRNLEQLGTVNPGSGTSSRLREPE 464

RESULT 19  
US-09-334-476-1  
; Sequence 1, Application US/09334476  
; Patent No. 6162901  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Lal, Preeti  
; TITLE OF INVENTION: NEW VESICLE TRANSPORT PROTEIN  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Dr.  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/334,476  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/191,279  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0350 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166

```
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 593 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; IMMEDIATE SOURCE:
;   LIBRARY: MMLR20T01
;   CLONE: 475485
; US-09-334-476-1

Query Match      24.0%; Score 50; DB 3; Length 593;
Best Local Similarity 31.1%; Pred. No. 27;
Matches 14; Conservative 9; Mismatches 12; Indels 10; Gaps 2;

QY 4 NLRILNKAALRKHAHTSMVRNFY-----GKPVQSOLKPRD 39
Db 421 NLAKLIQHNV-QAHSSLRINLQGLGTVTPNGSGTSRLLEPR 464

RESULT 20
US-09-543-681A-4698
; Sequence 4698, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 4698
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Proteus mirabilis
; US-09-543-681A-4698

Query Match      23.8%; Score 49.5; DB 4; Length 475;
Best Local Similarity 42.3%; Pred. No. 25;
Matches 11; Conservative 6; Mismatches 4; Indels 5; Gaps 1;

QY 9 LNKAA-----LRKAHTSMVRNFYK 29
Db 302 LNQGADVNDMRQALQALRSFYK 327

RESULT 21
US-09-252-991A-25705
; Sequence 25705, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25705
; LENGTH: 547
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-25705

Query Match      23.6%; Score 49; DB 4; Length 547;
Best Local Similarity 27.1%; Pred. No. 35;

US-09-252-991A-29653
; Sequence 29653, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29653
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-29653

Query Match      23.3%; Score 48.5; DB 4; Length 274;
Best Local Similarity 46.2%; Pred. No. 18;
Matches 12; Conservative 4; Mismatches 9; Indels 1; Gaps 1;

QY 12 AALRKAHTSMVRNFYKPVQSOLK 37
Db 179 AALRTQVDALRNFOQG-FVQAAEQ 203

RESULT 22
US-09-489-039A-7879
; Sequence 7879, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 7879
; LENGTH: 550
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
; US-09-489-039A-7879

Query Match      23.6%; Score 49; DB 4; Length 550;
Best Local Similarity 48.0%; Pred. No. 36;
Matches 12; Conservative 2; Mismatches 7; Indels 4; Gaps 1;

QY 16 KAHTSMVRNFYKPVQSOLKPRDL 40
Db 371 KYHT-----GFRHAMPVRQQLKTRTL 391

RESULT 23
US-09-252-991A-29653
; Sequence 29653, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29653
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-29653

Query Match      23.3%; Score 48.5; DB 4; Length 274;
Best Local Similarity 46.2%; Pred. No. 18;
Matches 12; Conservative 4; Mismatches 9; Indels 1; Gaps 1;

QY 12 AALRKAHTSMVRNFYKPVQSOLK 37
Db 179 AALRTQVDALRNFOQG-FVQAAEQ 203

RESULT 24
US-08-428-488-13
; Sequence 13, Application US/08428488
; Patent No. 5624894
; GENERAL INFORMATION:
; APPLICANT: BODOR, Nicholas S.
```

;/ TITLE OF INVENTION: BRAIN-ENHANCED DELIVERY OF NEUROACTIVE  
;/ TITLE OF INVENTION: PEPTIDES BY SEQUENTIAL METABOLISM  
;/ NUMBER OF SEQUENCES: 107  
;/ CORRESPONDENCE ADDRESS:  
;/ ADDRESSEE: Burns, Doane, Swecker & Mathis  
;/ STREET: P.O. Box 1404  
;/ CITY: Alexandria  
;/ STATE: Virginia  
;/ COUNTRY: United States  
;/ ZIP: 22313-1404  
;/ COMPUTER READABLE FORM:  
;/ MEDIUM TYPE: Floppy disk  
;/ COMPUTER: IBM PC compatible  
;/ OPERATING SYSTEM: PC-DOS/MS-DOS  
;/ SOFTWARE: PatentIn Release #1.0, Version #1.25  
;/ CURRENT APPLICATION DATA:  
;/ APPLICATION NUMBER: US/08/428,488  
;/ FILING DATE: 27-APR-1995  
;/ CLASSIFICATION: 514  
;/ ATTORNEY/AGENT INFORMATION:  
;/ NAME: Baumeister, Mary Katherine  
;/ REGISTRATION NUMBER: 26,254  
;/ REFERENCE/DOCKET NUMBER: 028724-087  
;/ TELECOMMUNICATION INFORMATION:  
;/ TELEPHONE: (703) 836-6620  
;/ TELEFAX: (703) 836-2021  
;/ INFORMATION FOR SEQ ID NO: 13:  
;/ SEQUENCE CHARACTERISTICS:  
;/ LENGTH: 39 amino acids  
;/ TYPE: amino acid  
;/ STRANDEDNESS: single  
;/ TOPOLOGY: linear  
;/ MOLECULE TYPE: peptide  
;/ FEATURE:  
;/ NAME/KEY: Modified-site  
;/ LOCATION: 1  
;/ OTHER INFORMATION: /note= "Position 1 = H-Ser."  
;/ FEATURE:  
;/ NAME/KEY: Modified-site  
;/ LOCATION: 39  
;/ OTHER INFORMATION: /note= "Position 39 = Phe-OH."  
;/ FEATURE:  
;/ NAME/KEY: Modified-site  
;/ LOCATION: 30  
;/ OTHER INFORMATION: /note= "Position 30 = Glu-NH2."  
;/ US-08-428-488-13

Query Match 23.1%; Score 48; DB 1; Length 39;  
Best Local Similarity 50.0%; Pred. No. 2.1;  
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 22 VRNFRYKPKVQSOLKP 37  
Db 4 MEHFRYKPKVKKRRP 19

RESULT 25  
US-09-732-210-98  
;/ Sequence 98, Application US/09732210  
;/ Patent No. 6573361  
;/ GENERAL INFORMATION:  
;/ APPLICANT: Bunkers, Greg J.  
;/ APPLICANT: Liang, Jihong  
;/ APPLICANT: Mittanck, Cindy A.  
;/ APPLICANT: Seale, Jeffrey W.  
;/ APPLICANT: Wu, Yonnie S.  
;/ TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use  
;/ FILE REFERENCE: 38-21(15036)B  
;/ CURRENT APPLICATION NUMBER: US/09/732,210  
;/ CURRENT FILING DATE: 2000-12-07  
;/ PRIOR APPLICATION NUMBER: US 60/169,513  
;/ PRIOR FILING DATE: 1999-12-07  
;/ PRIOR APPLICATION NUMBER: US 60/169,340

;/ PRIOR FILING DATE: 1999-12-07  
;/ NUMBER OF SEQ ID NOS: 1753  
;/ SEQ ID NO 98  
;/ LENGTH: 89  
;/ TYPE: PRT  
;/ ORGANISM: Schizosaccharomyces pombe  
;/ US-09-732-210-98

Query Match 23.1%; Score 48; DB 4; Length 89;  
Best Local Similarity 45.0%; Pred. No. 5.8;  
Matches 9; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

Qy 14 LRKAHTSMVRNFRYKPKVQS 33  
Db 67 LKKVHRSFKNGFRAGKPTSA 86

RESULT 26  
US-09-248-796A-19559  
;/ Sequence 19559, Application US/09248796A  
;/ Patent No. 6747137  
;/ GENERAL INFORMATION:  
;/ APPLICANT: Keith Weinstock et al  
;/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS  
;/ FILE REFERENCE: 107196.132  
;/ CURRENT APPLICATION NUMBER: US/09/248,796A  
;/ CURRENT FILING DATE: 1999-02-12  
;/ PRIOR APPLICATION NUMBER: US 60/074,725  
;/ PRIOR FILING DATE: 1998-02-13  
;/ PRIOR APPLICATION NUMBER: US 60/096,409  
;/ PRIOR FILING DATE: 1998-08-13  
;/ NUMBER OF SEQ ID NOS: 28208  
;/ SEQ ID NO 19559  
;/ LENGTH: 382  
;/ TYPE: PRT  
;/ ORGANISM: Candida albicans  
;/ US-09-248-796A-19559

Query Match 23.1%; Score 48; DB 4; Length 382;  
Best Local Similarity 40.7%; Pred. No. 33;  
Matches 11; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

Qy 14 LRKAHTSMVRNFRYKPKVQSOLKPRDL 40  
Db 167 LRAASTFIQRAFEYAQPEDQELLIKDL 193

RESULT 27  
US-09-188-811-2  
;/ Sequence 2, Application US/09188811  
;/ Patent No. 6037148  
;/ GENERAL INFORMATION:  
;/ APPLICANT: Khodadoust, Mehran  
;/ TITLE OF INVENTION: NOVEL MTX PROTEIN AND NUCLEIC ACID MOLECULES AND USES  
;/ FILE REFERENCE: MNI-046CP  
;/ CURRENT APPLICATION NUMBER: US/09/188,811  
;/ CURRENT FILING DATE: 1998-11-09  
;/ EARLIER APPLICATION NUMBER: 09/163,116  
;/ EARLIER FILING DATE: 1998-09-29  
;/ NUMBER OF SEQ ID NOS: 8  
;/ SOFTWARE: PatentIn Ver. 2.0  
;/ SEQ ID NO 2  
;/ LENGTH: 567  
;/ TYPE: PRT  
;/ ORGANISM: Homo sapiens  
;/ US-09-188-811-2

Query Match 23.1%; Score 48; DB 3; Length 567;  
Best Local Similarity 37.1%; Pred. No. 53;  
Matches 13; Conservative 5; Mismatches 9; Indels 8; Gaps 2;



TITLE OF INVENTION: DELIVERY OF NUCLEIC ACIDS TO CELLS

0.

;; APPLICANT: Nefsky, Bradley  
;; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto  
;; NUMBER OF SEQUENCES: 6  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
;; STREET: One Post Office Square  
;; CITY: Boston  
;; STATE: MA  
;; COUNTRY: USA  
;; ZIP: 02109-2170  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/392,163A  
;; FILING DATE:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/539,205  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Vincent, Matthew P.  
;; REGISTRATION NUMBER: 36,709  
;; REFERENCE/DOCKET NUMBER: CSV-005.01  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (617) 832-1000  
;; TELEFAX: (617) 832-7000  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 735 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-09-392-163A-2

Query Match 22.6%; Score 47; DB 4; Length 735;  
Best Local Similarity 45.0%; Pred. No. 1e+02;  
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 18 HTSMVNFYRGKPVQSQLK 37  
DB 534 HTFCVHNAYGEIIQHLKP 553

RESULT 37  
US-09-369-364A-9  
; Sequence 9, Application US/09369364A  
; Patent No. 6391610  
; GENERAL INFORMATION:  
; APPLICANT: Apte, Suneel  
; APPLICANT: Hurskainen, Tiina L.  
; APPLICANT: Hirohata, Satoshi  
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases  
; FILE REFERENCE: 26473/4007/10-30-00  
; CURRENT APPLICATION NUMBER: US/09/369,364A  
; CURRENT FILING DATE: 1999-08-06  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 9  
; LENGTH: 905  
; TYPE: PRT  
; ORGANISM: Mus musculus ADAMTS-8  
US-09-369-364A-9

Query Match 22.6%; Score 47; DB 3; Length 905;  
Best Local Similarity 41.7%; Pred. No. 1.3e+02;  
Matches 10; Conservative 2; Mismatches 12; Indels 0; Gaps 0;

QY 18 HTSMVNFYRGKPVQSQLKPRDLC 41  
DB 616 HTDLGDFLOWVPKYGVSPDRDC 639

RESULT 38  
US-09-428-711A-16  
; Sequence 16, Application US/09428711A  
; Patent No. 6358720  
; GENERAL INFORMATION:  
; APPLICANT: Muramatsu, Masaaki  
; APPLICANT: Shirasawa, Takuji  
; APPLICANT: Tokumitsu, Hiroshi  
; APPLICANT: No. 6358720uchi, Teruhisa  
; TITLE OF INVENTION: SERINE/THREONINE PROTEIN KINASE  
; FILE REFERENCE: 06501-045001  
; CURRENT APPLICATION NUMBER: US/09/428,711A  
; CURRENT FILING DATE: 1999-10-28  
; PRIOR APPLICATION NUMBER: PCT/JP98/01246  
; PRIOR FILING DATE: 1998-03-23  
; PRIOR APPLICATION NUMBER: JP 9/124798  
; PRIOR FILING DATE: 1997-04-28  
; NUMBER OF SEQ ID NOS: 21  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16  
; LENGTH: 1050  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-428-711A-16

Query Match 22.6%; Score 47; DB 3; Length 1050;  
Best Local Similarity 30.8%; Pred. No. 1.6e+02;  
Matches 12; Conservative 10; Mismatches 13; Indels 4; Gaps 1;

QY 7 ILLNKAALRKAKTSMVR----NFRYKPKVQSQLKPRDLC 41  
DB 144 ILLSNPAGRRANPNSIRVKIADFGFARYLQSNMMAATLC 182

RESULT 39  
US-09-803-286A-2  
; Sequence 2, Application US/09803286A  
; Patent No. 6664447  
; GENERAL INFORMATION:  
; APPLICANT: Tanksley, Steven D.  
; APPLICANT: Brommonschenkel, Sergio H.  
; TITLE OF INVENTION: PLANT GENE CONFERRING RESISTANCE TO TOSPOVIRUSES  
; FILE REFERENCE: 19603/3201  
; CURRENT APPLICATION NUMBER: US/09/803,286A  
; CURRENT FILING DATE: 2001-08-31  
; PRIOR APPLICATION NUMBER: 60/188,356  
; PRIOR FILING DATE: 2000-03-10  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 1247  
; TYPE: PRT  
; ORGANISM: Lycopersicon var.  
US-09-803-286A-2

Query Match 22.6%; Score 47; DB 4; Length 1247;  
Best Local Similarity 35.7%; Pred. No. 2e+02;  
Matches 15; Conservative 5; Mismatches 20; Indels 2; Gaps 1;

QY 2 LSNLRI--LLNKAALRKAKTSMVRNFRYKPKVQSQLKPRDLC 41  
DB 1101 LSGQLPSNLNKLVLGEIHIESVIPFIAGLPSLEYLQLQDVC 1142

RESULT 40  
US-09-413-814-42  
; Sequence 42, Application US/09413814  
; Patent No. 6225064  
; GENERAL INFORMATION:  
; APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH  
; APPLICANT: Bristol-Myers Squibb, Co.  
; APPLICANT: Beyer, Stefan



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 18, 2004, 02:50:12 ; Search time 95.1786 Seconds  
(without alignments)  
154.144 Million cell updates/sec

Title: US-08-765-244-1

Perfect score: 208

Sequence: 1 MLSNRLILNKAAALKAHTS.....VNFYRGKPVQSLKPRDLC 41

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1589859 seqs, 357834939 residues

Total number of hits satisfying chosen parameters: 1589859

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	208	100.0	41	US-08-765-244-1	Sequence 1, Appli
2	197	94.7	43	US-08-765-244-22	Sequence 22, Appli
3	92	44.2	32	US-10-061-395-48	Sequence 48, Appli
4	92	44.2	32	US-10-052-942-54	Sequence 54, Appli
5	58	27.9	99	US-09-855-604-639	Sequence 639, App
6	56	26.9	302	US-10-369-493-3761	Sequence 3761, Ap
7	56	26.9	1286	US-09-998-027-3	Sequence 3, Appli
8	56	26.9	1286	US-10-165-099-3	Sequence 3, Appli
9	55.5	26.7	761	US-10-270-875-27	Sequence 27, Appli
10	55.5	26.7	761	US-10-270-878-27	Sequence 27, Appli
11	55.5	26.7	761	US-10-270-786-27	Sequence 27, Appli
12	55.5	26.7	761	US-10-270-710-27	Sequence 27, Appli
13	55.5	26.7	761	US-10-270-859-27	Sequence 27, Appli

14	55.5	26.7	761	14	US-10-270-846-27	Sequence 27, Appli
15	55.5	26.7	1468	14	US-10-369-493-1976	Sequence 1976, Ap
16	53.5	25.7	383	14	US-10-369-493-17293	Sequence 17293, A
17	53	25.5	435	14	US-10-369-493-22531	Sequence 22531, A
18	52	25.0	112	17	US-10-425-115-326818	Sequence 326818, A
19	51.5	24.8	200	15	US-10-282-122A-57678	Sequence 57678, A
20	51	24.5	387	16	US-10-437-963-118548	Sequence 118548, A
21	50	24.0	72	17	US-10-425-115-212280	Sequence 212280, A
22	50	24.0	101	15	US-10-424-599-229228	Sequence 229228, A
23	50	24.0	142	15	US-10-108-260A-4711	Sequence 4711, Ap
24	50	24.0	193	15	US-10-264-237-2115	Sequence 2115, Ap
25	50	24.0	310	10	US-09-993-959-4	Sequence 4, Appli
26	50	24.0	310	17	US-10-825-632-3	Sequence 3, Appli
27	50	24.0	465	17	US-10-825-632-5	Sequence 5, Appli
28	50	24.0	590	15	US-10-205-331-78	Sequence 78, Appli
29	50	24.0	882	9	US-09-976-674-1	Sequence 1, Appli
30	50	24.0	882	14	US-10-054-776-2	Sequence 2, Appli
31	50	24.0	882	14	US-10-170-789-38	Sequence 38, Appli
32	50	24.0	882	15	US-10-311-035-9	Sequence 9, Appli
33	50	24.0	882	15	US-10-072-012-622	Sequence 622, App
34	50	24.0	882	15	US-10-415-122-6	Sequence 6, Appli
35	50	24.0	882	17	US-10-825-632-1	Sequence 1, Appli
36	50	24.0	883	15	US-10-072-012-621	Sequence 621, App
37	49	23.6	81	15	US-10-424-599-158049	Sequence 158049, A
38	49	23.6	108	17	US-10-425-115-197217	Sequence 197217, A
39	49	23.6	129	15	US-10-424-599-281450	Sequence 281450, A
40	49	23.6	137	15	US-10-424-599-252374	Sequence 252374, A
41	49	23.6	193	17	US-10-739-930-8796	Sequence 8796, Ap
42	49	23.6	201	15	US-10-282-122A-55605	Sequence 55605, A
43	49	23.6	268	9	US-09-739-907-114	Sequence 114, App
44	49	23.6	268	11	US-09-938-671-114	Sequence 114, App
45	49	23.6	355	14	US-10-369-493-21409	Sequence 21409, A

#### ALIGNMENTS

#### RESULT 1

US-08-765-244-1  
; Sequence 1, Application US/08765244  
; Publication No. US20010008771A1  
; GENERAL INFORMATION:  
; APPLICANT: Seibel, Peter  
; APPLICANT: Seibel, Andrea  
; TITLE OF INVENTION: CHIMICAL PEPTIDE-NUCLEIC ACID  
; TITLE OF INVENTION: FRAGMENT, PROCESS FOR PRODUCING THE SAME AND ITS USES FOR  
; TITLE OF INVENTION: APPROPRIATELY INTRODUCING NUCLEIC ACIDS INTO CELL ORGANELLES  
; TITLE OF INVENTION: AND CELLS  
; FILE REFERENCE: 8484-0018-999  
; CURRENT APPLICATION NUMBER: US/08/765,244  
; CURRENT FILING DATE: 1997-10-30  
; PRIOR APPLICATION NUMBER: PCT/DE95/00775  
; PRIOR FILING DATE: 1995-06-11  
; PRIOR APPLICATION NUMBER: DE P 44 21 079.5  
; PRIOR FILING DATE: 1994-06-16  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 41  
; TYPE: PRT  
; ORGANISM: Rattus rattus  
; FEATURE:  
US-08-765-244-1

Query Match 100.0%; Score 208; DB 8; Length 41;  
Best Local Similarity 100.0%; Pred. No. 5,1e-23;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLSNRLILNKAAALKAHTSMVNFYRGKPVQSLKPRDLC 41  
Db 1 MLSNRLILNKAAALKAHTSMVNFYRGKPVQSLKPRDLC 41



```
; PRIOR FILING DATE: 1997-09-11
; NUMBER OF SEQ ID NOS: 935
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 639
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-855-604-639

Query Match      27.9%; Score 58; DB 12; Length 99;
Best Local Similarity 39.4%; Pred. No. 1.7;
Matches 13; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

QY 6 RLLNKAALRKAHTSMVRNFRYKPVQSOLK 38
Db 53 RVILCAHTKANKQSRARTLRPLRLRIALRPR 85

RESULT 6
US-10-369-493-3761
; Sequence 3761, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 3761
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Neurospora crassa
US-10-369-493-3761

Query Match      26.9%; Score 56; DB 14; Length 302;
Best Local Similarity 64.7%; Pred. No. 12;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 12 AALRKAHTSMVRNFRYK 28
Db 271 AAMTKAHVSQMRFRYK 287

RESULT 7
US-09-998-027-3
; Sequence 3, Application US/09998027
; Publication No. US20030093819A1
; GENERAL INFORMATION:
; APPLICANT: D'Andrea et al.
; TITLE OF INVENTION: Methods and Compositions for the
; TITLE OF INVENTION: Diagnosis and Treatment of Cancers Associated with Defective
; TITLE OF INVENTION: DNA Repair Mechanisms
; FILE REFERENCE: 2486/101
; CURRENT APPLICATION NUMBER: US/09/998,027
; CURRENT FILING DATE: 2001-11-02
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1286
; TYPE: PRT
; ORGANISM: A. thaliana
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)...(1286)
; OTHER INFORMATION: Plantfancd2
```

```
US-09-998-027-3

Query Match      26.9%; Score 56; DB 10; Length 1286;
Best Local Similarity 35.3%; Pred. No. 64;
Matches 12; Conservative 8; Mismatches 14; Indels 0; Gaps 0;

QY 3 SNLRILLNKAALRKAHTSMVRNFRYKPVQSOLK 36
Db 107 SNLRMLSSSTTKRDESLVRNLLLVSPQLDQ 140

RESULT 8
US-10-165-099-3
; Sequence 3, Application US/10165099
; Publication No. US20030188326A1
; GENERAL INFORMATION:
; APPLICANT: D'Andrea, Alan
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS OF CANCER SUSCEPTIBILITY
; TITLE OF INVENTION: DEFECTIVE DNA REPAIR MECHANISMS AND TREATMENT THEREOF
; FILE REFERENCE: 7032/2055
; CURRENT APPLICATION NUMBER: US/10/165,099
; CURRENT FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: US 09/998,027
; PRIOR FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: US 60/245,756
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 352
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1286
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-165-099-3

Query Match      26.9%; Score 56; DB 14; Length 1286;
Best Local Similarity 35.3%; Pred. No. 64;
Matches 12; Conservative 8; Mismatches 14; Indels 0; Gaps 0;

QY 3 SNLRILLNKAALRKAHTSMVRNFRYKPVQSOLK 36
Db 107 SNLRMLSSSTTKRDESLVRNLLLVSPQLDQ 140

RESULT 9
US-10-270-875-27
; Sequence 27, Application US/10270875
; Publication No. US20030082741A1
; GENERAL INFORMATION:
; APPLICANT: Sigridur Hjorleifsdottir
; APPLICANT: Gudmundur O. Hreggvidsson
; APPLICANT: Olafur H. Fridjonsson
; APPLICANT: Arnthor Aevarsson
; APPLICANT: Jakob K. Kristjansson
; TITLE OF INVENTION: Bacteriophage RM378 of a Thermophilic
; TITLE OF INVENTION: Host Organism
; FILE REFERENCE: 2739.1001-001
; CURRENT APPLICATION NUMBER: US/10/270,875
; CURRENT FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: US/09/585,858
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/137,120
; PRIOR FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 761
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-270-875-27

Query Match      26.7%; Score 55.5; DB 14; Length 761;
Best Local Similarity 29.4%; Pred. No. 41;
Matches 15; Conservative 4; Mismatches 21; Indels 11; Gaps 1;
```

Db 62 MPNLRCLSLSIOTLMNPKENKQEI V SITLSAYRNISLDSPIENIKPDDL 112

Query Match 26.7%; Score 55.5; DB 14; Length 761;  
Best Local Similarity 29.4%; Pred. No. 41;  
Matches 15; Conservative 4; Mismatches 21; Indels 11

62 MPNLRCLSLSIOTLMNPKENKOEIVSITLSAYRNISLDSPINIKPDDLC 112

Query Match	26.7%	Score 55.5;	DB 14;	Length 761;
Best Local Similarity	29.4%;	Pred. NO. 41;		
Matches 15: Conservative	4;	Mismatches 21;	Indels 11	

[illegible]

RESULT 12  
US-10-270-710-27  
; Sequence 27, Application US/10270710  
; Publication No. US20030092128A1

US-10-270-710-27  
; Sequence 27, Application US/10270710  
; Publication No. US20030092128A1

Query Match	26.7%;	Score 55.5;	DB 14;	Length 761;
Best Local Similarity	29.4%;	Pred. No. 41;		
Matches 15;	Conservative	4;	Mismatches	21;
			Indels	11;
			Gaps	1;

Db 62 MPNLRCLSLSIOTLMNPKENKQEI VSI TLSAYRNISLSDSPIPENIKPDDLC 11

Query Match	26.7%	Score 55.5;	DB 14;	Length 761;
Best Local Similarity	29.4%;	Pred. No. 41;		
Matches 15;	Conservative	4;	Mismatches	21;
				Indels 11;
				Gaps 1;

db 62 MPNLRCLSLSIOTLMNPKENKQEI VSI TLSAYRNISLSDSPIPENIKPDDL 112



```

; Sequence 17293, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCES: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 17293
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Bacillus halodurans
; US-10-369-493-17293

Query Match 25.7%; Score 53.5; DB 14; Length 383;
Best Local Similarity 42.9%; Pred. No. 37;
Matches 15; Conservative 3; Mismatches 14; Indels 3; Gaps 1;

QY 2 LSNRLILLNKAALRKAHTSMVNFYRGKPVOSQLK 36
: ||||| ||| ||| ||| :
Db 114 IQNKILLNSAATEKATGSPARG--GKPTTATK 145

RESULT 17
US-10-369-493-22531
; Sequence 22531, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCES: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 22531
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; US-10-369-493-22531

Query Match 25.5%; Score 53; DB 14; Length 435;
Best Local Similarity 46.4%; Pred. No. 50;
Matches 13; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 3 SNIRILLNKAALRKAHTSMVNFYRGK 30
||| ||| ||| :
Db 28 SNLKILQNKRALSKNDSSSKQVQDSKP 55

RESULT 18
US-10-425-115-326818
; Sequence 326818, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei

```





```
; PRIOR APPLICATION NUMBER: AU P05709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU P02762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-825-632-3

Query Match      24.0%; Score 50; DB 17; Length 310;
Best Local Similarity 44.0%; Pred. No. 94;
Matches 11; Conservative 6; Mismatches 6; Indels 2; Gaps 1;

QY 17 AHTSMVRNF--RYGKPVQSQLKPRD 39
Db 251 AHTSILLFLVRAGKPYDLQIYPQE 275

RESULT 27
US-10-825-632-5
; Sequence 5, Application US/10825632
; Publication No. US20040191826A1
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: GORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: FCSB-100-DIV. 1
; CURRENT APPLICATION NUMBER: US/10/825,632
; CURRENT FILING DATE: 2004-04-15
; PRIOR APPLICATION NUMBER: US 10/070,464
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU P05709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU P02762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-825-632-5

Query Match      24.0%; Score 50; DB 17; Length 465;
Best Local Similarity 44.0%; Pred. No. 1.5e+02;
Matches 11; Conservative 6; Mismatches 6; Indels 2; Gaps 1;

QY 17 AHTSMVRNF--RYGKPVQSQLKPRD 39
Db 406 AHTSILLFLVRAGKPYDLQIYPQE 430

RESULT 28
US-10-205-331-78
; Sequence 78, Application US/10205331
; Publication No. US20040058326A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; APPLICANT: Lee, Kevin
; APPLICANT: Dixon, Alistair
; APPLICANT: Brooksbank, Robert
; APPLICANT: Pinnock, Robert
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
; FILE REFERENCE: WL-A-018199
; CURRENT APPLICATION NUMBER: US/10/205,331
; CURRENT FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: GB 0118354.0
; PRIOR FILING DATE: 2001-07-27

; PRIOR APPLICATION NUMBER: AU P05709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU P02762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 78
; LENGTH: 590
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-205-331-78

Query Match      24.0%; Score 50; DB 15; Length 590;
Best Local Similarity 31.1%; Pred. No. 2e+02;
Matches 14; Conservative 9; Mismatches 12; Indels 10; Gaps 2;

QY 4 NLRILNKAAALRKAAHTSMVRNFRY-----GKPVQSQLKPRD 39
Db 418 NLAKLIQHNV-QAHSSLIRNLEQLGGTVPFGSGTSSRLPRE 461

RESULT 29
US-09-976-674-1
; Sequence 1, Application US/09976674
; Patent No. US20020115843A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-1

Query Match      24.0%; Score 50; DB 9; Length 882;
Best Local Similarity 44.0%; Pred. No. 3.2e+02;
Matches 11; Conservative 6; Mismatches 6; Indels 2; Gaps 1;

QY 17 AHTSMVRNF--RYGKPVQSQLKPRD 39
Db 823 AHTSILLFLVRAGKPYDLQIYPQE 847

RESULT 30
US-10-054-776-2
; Sequence 2, Application US/10054776
; Publication No. US20030165818A1
; GENERAL INFORMATION:
; APPLICANT: Mark Robert Edbrooke
; APPLICANT: Alan Peter Lewis
; TITLE OF INVENTION: NOVEL PROTEIN
; FILE REFERENCE: QG10420S
; CURRENT APPLICATION NUMBER: US/10/054,776
; CURRENT FILING DATE: 2002-01-23
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-054-776-2

Query Match      24.0%; Score 50; DB 14; Length 882;
Best Local Similarity 44.0%; Pred. No. 3.2e+02;
Matches 11; Conservative 6; Mismatches 6; Indels 2; Gaps 1;
```

```
QY 17 AHTSMVRNF--RYGKPVQSOLKPRD 39
      |||:::| | ||| |:::|
Db 823 AHTSILLSFLVRAGKPYDLQIYPQE 847

RESULT 31
US-10-170-789-38
; Sequence 38, Application US/10170789
; Publication No. US20030180930A1
; GENERAL INFORMATION:
; APPLICANT: Rachel E. Meyers
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Williamson, Mark
; APPLICANT: Weich, Nadine
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,
; FILE REFERENCE: 10448-191001
; CURRENT APPLICATION NUMBER: US/10/170,789
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06525
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/882,166
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: US 09/861,801
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: PCT/US01/16549
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 60/205,508
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/801,267
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: PCT/US01/07138
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/187,454
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/829,671
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: PCT/US01/40483
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: US 60/197,508
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/961,721
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/29904
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 60/235,023
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 10/045,367
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 60/246,561
; PRIOR FILING DATE: 2000-11-07
; PRIOR APPLICATION NUMBER: US 09/801,275
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: PCT/US01/07074
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/187,420
```

```
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-789-38

Query Match 24.0%; Score 50; DB 14; Length 882;
Best Local Similarity 44.0%; Pred. No. 3.2e+02;
Matches 11; Conservative 6; Mismatches 6; Indels 2; Gaps 1;

QY 17 AHTSMVRNF--RYGKPVQSOLKPRD 39
      |||:::| | ||| |:::|
Db 823 AHTSILLSFLVRAGKPYDLQIYPQE 847

RESULT 32
US-10-311-035-9
; Sequence 9, Application US/10311035
; Publication No. US20040023243A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: ELLIOTT, Vicki
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: LAL, Preeti
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: DELGEANE, Angelo M.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: NGUYEN, Dannie B.
; APPLICANT: LEE, Ernestine A.
; APPLICANT: HAFALIA, Adril
; APPLICANT: KHAN, Farrah A.
; APPLICANT: CHAWLA, Narinder K.
; APPLICANT: YAO, Monique G.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: ARVIZU, Chandra S.
; APPLICANT: TANG, Y. Tom
; APPLICANT: WALSH, Roderick T.
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: LU, Yan
; APPLICANT: RAMKUMAR, Jayalaximi
; APPLICANT: XU, Yuming
; APPLICANT: REDDY, Roopa
; APPLICANT: DAS, Depopriya
; APPLICANT: KEARNEY, Liam
; APPLICANT: KALLICK, Deborah A.
; TITLE OF INVENTION: Proteases
; FILE REFERENCE: PI-0123 PCT
; CURRENT APPLICATION NUMBER: US/10/311,035
; PRIOR FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/212,336; 60/213,955; 60/215,396; 60/216,821; 60/218,946
; PRIOR FILING DATE: 2000-06-16; 2000-06-22; 2000-06-29; 2000-07-07; 2000-07-14
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 9
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040023243A1 7160544CD1
US-10-311-035-9

Query Match 24.0%; Score 50; DB 15; Length 882;
Best Local Similarity 44.0%; Pred. No. 3.2e+02;
Matches 11; Conservative 6; Mismatches 6; Indels 2; Gaps 1;

QY 17 AHTSMVRNF--RYGKPVQSOLKPRD 39
      |||:::| | ||| |:::|
```

Db 823 AHTSILLSFLVRACKPYDLQIYPQE 847

## RESULT 33

US-10-072-012-622

; Sequence 622, Application US/10072012

; Publication No. US20040033493A1

; GENERAL INFORMATION:

; APPLICANT: Tchernev, Velizar

; APPLICANT: Spytek, Kimberly

; APPLICANT: Zerhusen, Bryan

; APPLICANT: Patturajan, Meera

; APPLICANT: Shimkets, Richard

; APPLICANT: Li, Li

; APPLICANT: Gangolli, Esha

; APPLICANT: Padigar, Muralidhara

; APPLICANT: Anderson, David W.

; APPLICANT: Rastelli, Luca

; APPLICANT: Miller, Charles E.

; APPLICANT: Gerlach, Valerie

; APPLICANT: Taupier Jr, Raymond J.

; APPLICANT: Gusev, Vladimir Y.

; APPLICANT: Colman, Steven D.

; APPLICANT: Wolenc, Adam R.

; APPLICANT: Pena, Carol E. A

; APPLICANT: Furtak, Katarzyna

; APPLICANT: Grosse, William M.

; APPLICANT: Alsobrook II, John P.

; APPLICANT: Lepley, Denise M.

; APPLICANT: Rieger, Daniel K.

; APPLICANT: Burgess, Catherine E.

; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same

; FILE REFERENCE: 21402-258

; CURRENT APPLICATION NUMBER: US/10/072,012

; CURRENT FILING DATE: 2002-01-31

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: 60/265,102

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: 60/265,514

; PRIOR FILING DATE: 2001-01-31

; PRIOR APPLICATION NUMBER: 60/265,517

; PRIOR FILING DATE: 2001-01-31

; PRIOR APPLICATION NUMBER: 60/265,412

; PRIOR FILING DATE: 2001-01-31

; PRIOR APPLICATION NUMBER: 60/265,395

; PRIOR FILING DATE: 2001-01-31

; PRIOR APPLICATION NUMBER: 60/266,406

; PRIOR FILING DATE: 2001-02-02

; PRIOR APPLICATION NUMBER: 60/266,767

; PRIOR FILING DATE: 2001-02-05

; PRIOR APPLICATION NUMBER: 60/267,057

; PRIOR FILING DATE: 2001-02-07

; PRIOR APPLICATION NUMBER: 60/266,975

; PRIOR FILING DATE: 2001-02-07

; PRIOR APPLICATION NUMBER: 60/267,459

; PRIOR FILING DATE: 2001-02-08

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 1391

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 622

; LENGTH: 882

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-072-012-622

Query Match 24.0%; Score 50; DB 15; Length 882;

Best Local Similarity 44.0%; Pred. No. 3.2e+02;

Matches 11; Conservative 6; Mismatches 6; Indels 2; Gaps 1;

QY 17 AHTSMVRNF--RYGKPVQSQLKPRD 39

|||||: :| | | | | :| :| :

Db 823 AHTSILLSFLVRACKPYDLQIYPQE 847

## RESULT 34

US-10-415-122-6

; Sequence 6, Application US/10415122

; Publication No. US20040053369A1

; GENERAL INFORMATION:

; APPLICANT: THE UNIVERSITY OF SYDNEY

; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES

; FILE REFERENCE: FP15217

; CURRENT APPLICATION NUMBER: US/10/415,122

; CURRENT FILING DATE: 2003-08-07

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 6

; LENGTH: 882

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-415-122-6

Query Match 24.0%; Score 50; DB 15; Length 882;

Best Local Similarity 44.0%; Pred. No. 3.2e+02;

Matches 11; Conservative 6; Mismatches 6; Indels 2; Gaps 1;

QY 17 AHTSMVRNF--RYGKPVQSQLKPRD 39

|||||: :| | | | | :| :| :

Db 823 AHTSILLSFLVRACKPYDLQIYPQE 847

## RESULT 35

US-10-825-632-1

; Sequence 1, Application US/10825632

; Publication No. US20040191826A1

; GENERAL INFORMATION:

; APPLICANT: ABBOTT, Catherine Anne

; APPLICANT: GORRELL, Mark Douglas

; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES

; FILE REFERENCE: FCSB-100-Div. 1

; CURRENT APPLICATION NUMBER: US/10/825,632

; CURRENT FILING DATE: 2004-04-15

; PRIOR APPLICATION NUMBER: US 10/070,464

; PRIOR FILING DATE: 2002-07-18

; PRIOR APPLICATION NUMBER: PCT/AU00/01085

; PRIOR FILING DATE: 2000-09-11

; PRIOR APPLICATION NUMBER: AU PQ5709

; PRIOR FILING DATE: 2000-02-18

; PRIOR APPLICATION NUMBER: AU PQ2762

; PRIOR FILING DATE: 1999-09-10

; NUMBER OF SEQ ID NOS: 31

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 882

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-10-825-632-1

Query Match 24.0%; Score 50; DB 17; Length 882;

Best Local Similarity 44.0%; Pred. No. 3.2e+02;

Matches 11; Conservative 6; Mismatches 6; Indels 2; Gaps 1;

QY 17 AHTSMVRNF--RYGKPVQSQLKPRD 39

|||||: :| | | | | :| :| :

Db 823 AHTSILLSFLVRACKPYDLQIYPQE 847

## RESULT 36

US-10-072-012-621

; Sequence 621, Application US/10072012

; Publication No. US20040033493A1

; GENERAL INFORMATION:

; APPLICANT: Tchernev, Velizar

; APPLICANT: Spytek, Kimberly

; APPLICANT: Zerhusen, Bryan

; APPLICANT: Patturajan, Meera

; APPLICANT: Shimkets, Richard



Query Match 23.6%; Score 49; DB 15; Length 129;  
Best Local Similarity 26.0%; Pred. No. 48;  
Matches 13; Conservative 10; Mismatches 15; Indels 12; Gaps 1;  
Qy 2 LSNRLILNKAALRKHAHTSMVRNFRYK-----PVOSQLKPRD 39  
Db 34 LKNLLVLSNPAPIFKGNPLTKSPQKSDIHFQKTLTQPLPKYKPKN 83

RESULT 40  
US-10-424-599-252374  
; Sequence 252374, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 252374  
; LENGTH: 137  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_69922C.1.pep  
US-10-424-599-252374

Query Match 23.6%; Score 49; DB 15; Length 137;  
Best Local Similarity 30.2%; Pred. No. 51;  
Matches 16; Conservative 6; Mismatches 13; Indels 18; Gaps 2;  
Qy 1 MLSNL---RILLNKAALRKHAHT-----SMVRNFRYKPKVQSOL 35  
Db 27 MLQKLCCKHLLREITVKKAHSIIITKWTFRTHRYTINMWLQFYAAPYNCL 79

Search completed: December 18, 2004, 03:07:05  
Job time : 97.1786 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 18, 2004, 02:37:05 ; Search time 22.9405 Seconds  
(without alignments)  
171.962 Million cell updates/sec

Title: US-08-765-244-1  
Perfect score: 208  
Sequence: 1 MSLNRLILLNKAALRKAHTS.....VNFYGRKPVQSQLKPRDLC 41

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: Pir1:\*  
2: Pir2:\*  
3: Pir3:\*  
4: Pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	179	86.1	354	1 OWRT	ornithine carbamoy
2	144	69.2	354	1 OWMS	ornithine carbamoy
3	125	60.1	354	1 OWHU	ornithine carbamoy
4	112	53.8	41	1 I52779	ornithine transcar
5	57	27.4	350	2 A48421	ornithine transcar
6	57	27.4	351	2 F90409	GTP binding conser
7	56	26.9	1286	2 B71413	hypothetical prote
8	55.5	26.7	1468	2 S58250	DNA-directed DNA p
9	54.5	26.2	138	2 F75040	hypothetical prote
10	53.5	25.7	260	2 G70233	hypothetical prote
11	53.5	25.7	383	2 B83922	short-chain-specif
12	53	25.5	435	2 S31290	cyclin B5 - yeast
13	52.5	25.2	185	2 B81708	translation elonga
14	51.5	24.8	306	2 T52305	En/Spm-like transp
15	51.5	24.8	314	2 F75504	probable pilin, ty
16	51	24.5	240	2 B45359	proopiomelanocorti
17	51	24.5	258	1 CTGNPK	corticotropin / li
18	51	24.5	506	2 T50211	WD-repeat protein
19	51	24.5	1779	2 T23130	hypothetical prote
20	49.5	23.8	181	2 T32637	hypothetical prote
21	49.5	23.8	425	2 I40646	sensor-like protei
22	49	23.6	178	2 H86026	hypothetical prote
23	49	23.6	178	2 F91180	hypothetical prote
24	49	23.6	276	2 AD3351	spoU family of rRN
25	49	23.6	300	2 T32702	hypothetical prote
26	49	23.6	320	2 AB0194	pseudouridylylate sy
27	49	23.6	353	2 T25678	hypothetical prote
28	49	23.6	355	1 B69518	GTP-binding protei
29	49	23.6	531	1 NNEC2	anthranilate synth

30	49	23.6	531	2 AE0653	anthranilate synth
31	49	23.6	531	2 E85761	hypothetical prote
32	49	23.6	531	2 C90858	anthranilate synth
33	49	23.6	538	2 F83354	probable sulfatase
34	49	23.6	986	1 A29646	invasin - Yersinia
35	48.5	23.3	333	2 A10802	Div protein [import
36	48.5	23.3	945	2 S77052	cation-transportin
37	48.5	23.3	1114	1 S05582	protein-tyrosine k
38	48	23.1	89	2 T43306	ribosomal protein
39	48	23.1	260	2 B75181	proteasome endopep
40	48	23.1	260	2 A71033	probable proteasom
41	48	23.1	362	2 B87613	hypothetical prote
42	48	23.1	633	2 S76749	hypothetical prote
43	48	23.1	765	2 T04186	subtilisin-like pr
44	48	23.1	893	2 T29621	hypothetical prote
45	47.5	22.8	260	2 T11798	cytochrome-c oxida

## ALIGNMENTS

## RESULT 1

OWRT

ornithine carbamoyltransferase (EC 2.1.3.3) precursor - rat

N:Alternate names: citrulline phosphorylase; ornithine transcarbamylase

C:Species: Rattus norvegicus (Norway rat)

C:Date: 28-Feb-1986 #sequence\_revision 28-Feb-1986 #text change 09-Jul-2004

C:Accession: A00563; A28042; A23090; S02466; I52976; I67609; I53457

R:Takiguchi, M.; Miura, S.; Mori, M.; Tatibana, M.; Nagata, S.; Kaziro, Y.

Proc. Natl. Acad. Sci. U.S.A. 81, 7412-7416, 1984

A:Title: Molecular cloning and nucleotide sequence of cDNA for rat ornithine carbamoyltransferase

A:Reference number: A00563; MUID:85063800; PMID:6095294

A:Accession: A00563

A:Molecule type: mRNA

A:Residues: 1-354 &lt;TAKI&gt;

A:Cross-references: UNIPROT:P00481; GB:K03040; NID:G205873; PIDN:AAA41768.1; PID:G205874

R:Takiguchi, M.; Murakami, T.; Miura, S.; Mori, M.

Proc. Natl. Acad. Sci. U.S.A. 84, 6136-6140, 1987

A:Title: Structure of the rat ornithine carbamoyltransferase gene, a large, X chromosome-

A:Reference number: A28042; MUID:87317609; PMID:3476935

A:Accession: A28042

A:Molecule type: DNA

A:Residues: 1-354 &lt;TAKE&gt;

A:Cross-references: GB:M16933; GB:J02957; NID:G205884; PIDN:AAA41769.1; PID:G205886

R:Kraus, J.P.; Hodges, P.E.; Williamson, C.L.; Horwich, A.L.; Kalousek, F.; Williams, K.F.

Nucleic Acids Res. 13, 943-952, 1985

A:Title: A cDNA clone for the precursor of rat mitochondrial ornithine transcarbamylase:

A:Reference number: A23090; MUID:85215524; PMID:3839075

A:Accession: A23090

A:Molecule type: mRNA

A:Residues: 1-38, 'P', 40-240, 'S', 242-354 &lt;KRA&gt;

A:Cross-references: GB:X01976

R:Aoki, Y.; Sunaga, H.; Suzuki, K.T.

Biochem. J. 250, 735-742, 1988

A:Title: A cadmium-binding protein in rat liver identified as ornithine carbamoyltransferase

A:Reference number: S02466; MUID:88268748; PMID:3390141

A:Accession: S02466

A:Molecule type: protein

A:Residues: 33-56; 293-302; 307-317; 322-329 &lt;AOK&gt;

R:McIntyre, P.; Graf, L.; Mercer, J.F.B.; Wake, S.A.; Hudson, P.J.; Hoogenraad, N.

DNA 4, 147-156, 1985

A:Title: The primary structure of the imported mitochondrial protein, ornithine transcarbamoyltransferase

A:Reference number: I52976; MUID:85203360; PMID:3838931

A:Accession: I52976

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-354 &lt;RES&gt;

A:Cross-references: GB:M11266; NID:G205871; PIDN:AAA41767.1; PID:G205872

R:McIntyre, P.; Graf, L.; Mercer, J.; Peterson, G.; Hudson, P.J.; Hoogenraad, N.

FEBS Lett. 177, 41-46, 1984

A:Title: A highly basic N-terminal extension of the mitochondrial matrix enzyme ornithine

A:Reference number: I53457; MUID:85051832; PMID:6548714

A:Accession: I67609

A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-102 <RE2>  
A;Cross-references: EMBL:X01178; NID:g56802; PIDN:CAA25618.1; PID:g56803  
A;Accession: I53457  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-43,'N',45-99,'R',101-102 <RE3>  
A;Cross-references: GB:K03041; NID:g205989; PIDN:AAA41771.1; PID:g205990  
C;Genetics:  
A;Introns: 26/2; 72/3; 100/1; 129/2; 180/3; 221/3; 239/3; 289/3; 335/3  
C;Superfamily: ornithine carbamoyltransferase; aspartate/ornithine carbamoyltransferase  
C;Keywords: arginine biosynthesis; homotrimer; mitochondrion; transferase; urea cycle  
F;1-32/Domain: transit peptide (mitochondrion) #status predicted <TNP>  
F;33-334/Product: ornithine carbamoyltransferase #status predicted <MAT>  
F;40-342/Domain: aspartate/ornithine carbamoyltransferase homology <ACT>  
  
Query Match 86.1%; Score 179; DB 1; Length 354;  
Best Local Similarity 92.9%; Pred. No. 1.7e-17;  
Matches 39; Conservative 0; Mismatches 1; Indels 2; Gaps 1;  
  
Qy 1 MLSNRLILNKALRKAHTSMVRNFRYKGPVQS--OLKPRDL 40  
Db 1 MLSNRLILNKALRKAHTSMVRNFRYKGPVQSQVLKGRDL 42  
  
RESULT 2  
OWNS  
ornithine carbamoyltransferase (EC 2.1.1.3.3) precursor - mouse  
N;Alternate names: citrulline phosphorylase; ornithine transcarbamylase  
C;Species: Mus musculus (house mouse)  
C;Date: 31-Mar-1993 #sequence revision 13-Mar-1997 #text\_change 09-Jul-2004  
C;Accession: A43609; S03407; I55252  
R;Veres, G.; Gibbs, R.A.; Scherer, S.E.; Caskey, C.T.  
Science 237, 415-417, 1987  
A;Title: The molecular basis of the sparse fur mouse mutation.  
A;Reference number: A43609; MUID:87263407; PMID:3603027  
A;Accession: A43609  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-354 <VER>  
A;Cross-references: UNIPROT:P11725; GB:M17030; NID:g200162; PIDN:AAA39865.1; PID:g200163  
R;Scherer, S.E.; Veres, G.; Caskey, C.T.  
Nucleic Acids Res. 16, 1593-1601, 1988  
A;Title: The genetic structure of mouse ornithine transcarbamylase.  
A;Reference number: S03407; MUID:88157717; PMID:2831503  
A;Accession: S03407  
A;Status: translation not shown  
A;Molecule type: DNA  
A;Residues: 1-194,'R',196-335 <SCH>  
A;Cross-references: EMBL:X07092  
R;Veres, G.; Craigen, W.J.; Caskey, C.T.  
J. Biol. Chem. 261, 7588-7591, 1986  
A;Title: The 5' flanking region of the ornithine transcarbamylase gene contains DNA sequence  
A;Reference number: I55252; MUID:86224037; PMID:3011788  
A;Accession: I55252  
A;Status: translation not shown; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-19,'LLWFDF' <RES>  
A;Cross-references: GB:M12716; NID:g200160; PIDN:AAA39864.1; PID:g554248  
A;Note: the end of this sequence is near the boundary of the cloned region and may be an  
C;Genetics:  
A;Gene: OTC  
A;Map position: X  
A;Introns: 26/2; 72/3; 100/1; 129/2; 180/3; 221/3; 239/3; 289/3  
C;Superfamily: ornithine carbamoyltransferase; aspartate/ornithine carbamoyltransferase  
C;Keywords: mitochondrion; transferase; urea cycle  
F;1-32/Domain: transit peptide (mitochondrion) #status predicted <TNP>  
F;33-335/Product: ornithine carbamoyltransferase #status predicted <MAT>  
F;40-342/Domain: aspartate/ornithine carbamoyltransferase homology <ACT>  
  
Query Match 69.2%; Score 144; DB 1; Length 354;  
Best Local Similarity 78.6%; Pred. No. 1.5e-12;

Matches 33; Conservative 2; Mismatches 5; Indels 2; Gaps 1;  
  
Qy 1 MLSNRLILNKALRKAHTSMVRNFRYKGPVQS--OLKPRDL 40  
Db 1 MLSNRLILNKALRKAHTSMVRNFRYKGPVQSQVLKGRDL 42  
  
RESULT 3  
OWHU  
ornithine carbamoyltransferase (EC 2.1.1.3.3) precursor - human  
N;Alternate names: citrulline phosphorylase; ornithine transcarbamylase  
C;Species: Homo sapiens (man)  
C;Date: 28-Feb-1986 #sequence revision 31-Mar-1993 #text\_change 09-Jul-2004  
C;Accession: A41444; A00562; I38078; JC4672; I59039; I54377  
R;Hata, A.; Tsuzuki, T.; Shimada, K.; Takiguchi, M.; Mori, M.; Matsuda, I.  
J. Biochem. 103, 302-308, 1988  
A;Title: Structure of the human ornithine transcarbamylase gene.  
A;Reference number: A41444; MUID:88227905; PMID:2836378  
A;Accession: A41444  
A;Molecule type: DNA  
A;Residues: 1-354 <HAT>  
A;Cross-references: UNIPROT:P00480; GB:D00230; NID:g219957; PIDN:BAA00161.1; PID:g219959  
R;Horwich, A.L.; Fenton, W.A.; Williams, K.R.; Kalousek, F.; Kraus, J.P.; Doolittle, R.F.  
Science 224, 1068-1074, 1984  
A;Title: Structure and expression of a complementary DNA for the nuclear coded precursor  
A;Reference number: A00562; MUID:84196410; PMID:6372096  
A;Accession: A00562  
A;Molecule type: mRNA  
A;Residues: 1-100,'P',102-110,'P',112-192,'CF',195-269,'R',271-354 <HOR>  
A;Cross-references: GB:D00230  
R;Hata, A.; Tsuzuki, T.; Shimada, K.; Takiguchi, M.; Mori, M.; Matsuda, I.  
J. Biochem. 100, 717-725, 1986  
A;Title: Isolation and characterization of the human ornithine transcarbamylase gene: at  
A;Reference number: I38078; MUID:87057134; PMID:3782067  
A;Accession: I38078  
A;Status: translation not shown  
A;Molecule type: DNA  
A;Residues: 1-26 <RES>  
A;Cross-references: EMBL:X04443; NID:g95162; PIDN:CAA28039.1; PID:g35163  
R;Wheeler, V.C.; Prodromou, C.P.; Pearl, L.H.; Williamson, R.; Coutelle, C.  
Gene 169, 251-255, 1996  
A;Title: Synthesis of a modified gene encoding human ornithine transcarbamylase for exper  
ment defect.  
A;Reference number: JC4672; MUID:96194812; PMID:8647457  
A;Accession: JC4672  
A;Molecule type: DNA  
A;Residues: 'M',33-100,'P',102-110,'P',112-192,'CF',195-269,'R',271-354 <WHE>  
A;Note: this report represents a synthetic gene designed for expression in (rather than t  
R;Horwich, A.L.; Kalousek, F.; Rosenberg, L.E.  
Proc. Natl. Acad. Sci. U.S.A. 82, 4930-4933, 1985  
A;Title: Arginine in the leader peptide is required for both import and proteolytic clea  
A;Reference number: I59039; MUID:85270440; PMID:3895227  
A;Accession: I59039  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-36 <RE2>  
A;Cross-references: GB:M1235; NID:g189408; PIDN:AAA59976.1; PID:g189409  
R;Gilbert-Dussardier, B.; Rabier, D.; Strautnieks, S.; Segues, B.; Bonnefont, J.P.; Munni  
Hum. Mol. Genet. 3, 831-832, 1994  
A;Title: A novel arginine (245) to glutamine change in exon 8 of the ornithine carbamoyl  
C pattern.  
A;Reference number: I54377; MUID:94362689; PMID:8081373  
A;Accession: I54377  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 269-276,'Q',278-289 <RE3>  
A;Cross-references: GB:S73640; NID:g688001; PIDN:AAB31859.1; PID:g688002  
A;Note: this sequence represents a disease defect in ornithine carbamoyltransferase  
C;Comment: The active enzyme is a dimer of identical chains with one tightly bound zinc

```

QY      1 MLSNRLINKAALRKHAHTSMVRNFYRYPKVPQSOLKPRDL 40
       ||:||:||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db      1 MLHHMRTIIN-ASWRYGNKCIVRFQFGSQ-TYSQLKGKRD 38

RESULT 6
F90409
GTP binding conserved hypothetical protein SSO2385 [imported] - Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C;Accession: F90409
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-y
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A;Description: Sulfolobus solfataricus complete genome.
A;Reference number: A99139
A;Accession: F90409
A>Status: preliminary
A:Molecule type: DNA
A;Residues: 1-351 <Kur>
A;Cross-references: UNIPROT:Q97W55; GB:AE006641; NID:g13815687; PIDN:AAK42533.1; GSPDB:B.
C;Genetics:
A;Gene: SSO2385

Query Match          27.4%; Score 57; DB 2; Length 351;
Best Local Similarity 32.7%; Pred. No. 3.3;
Matches 17; Conservative 8; Mismatches 9; Indels 18; Gaps 3;

QY      7 ILLNKAA-----LRKAHNTSMVNRFRY----CKPVQSQ-----LKPRDL 40
       ::||::||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db     294 LILKKGSTVLVDARKLHSLSLAENFRYARVMGSKVFQGKVGPSHILEDRI 345

RESULT 7
B71413
hypothetical protein dl3525w - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
A;Variety: columbia
C;Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 09-Jul-2004
C;Accession: B71413
R;Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirks
P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenegger, T.; Pohl, T.M.; Terryn, N.; Gle
avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
Nature 391, 485-488, 1998
A;Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech,
erhoft, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Anst
C.; Chludzisz, N.
A;Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thali
A;Reference number: A71400; MUID:98121113; PMID:9461215
A;Accession: B71413
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A;Residues: 1-1286 <BEV>
A;Cross-references: UNIPROT:O23351; GB:Z97337; NID:g2244829; PID:g2244854
C;Genetics:
A;Map position: ACOP9-4G3845
C;Superfamily: Arabidopsis thaliana hypothetical protein dl3525w

Query Match          26.9%; Score 56; DB 2; Length 1286;
Best Local Similarity 35.3%; Pred. No. 19;
Matches 12; Conservative 8; Mismatches 14; Indels 0; Gaps 0;

QY      3 SNLRILLNKAAALRKHAHTSMVRNFYRYPKVPQSOLK 36
       |||||::||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db     107 SNLRMLSSSTTKRDESLVERNLLLVSPILDIQ 140

RESULT 8
S58250
DNA-directed DNA polymerase (EC 2.7.7.7) - yeast (Saccharomyces cerevisiae)
N;Alternate names: DNA polymerase I; protein N2181; protein YNL102w
C;Species: Saccharomyces cerevisiae
```

C; Date: 13-Jan-1996 #sequence\_revision 23-Feb-1996 #text\_change 09-Jul-2004  
C; Accession: S58250; A30211; C42246; S31203; S63052; S22774  
R; Saiz, J.E.; Buित्रago, M.J.; Soler, A.; del Rey, F.; Revuelta, J.L.  
submitted to the EMBL Data Library, July 1995  
A; Description: The sequence of a 21.3 kb fragment from the left arm of yeast chromosome  
A; Reference number: S58246  
A; Accession: S58250  
A; Molecule type: DNA  
A; Residues: 1-1468 <SAI>  
A; Cross-references: UNIPROT:P13382; EMBL:Z50161; NID:G929846; PIDN:CAA90524.1; PID:G929851  
R; Pizagalli, A.; Valsasini, P.; Plevani, P.; Lucchini, G.  
Proc. Natl. Acad. Sci. U.S.A. 85, 3772-3776, 1988  
A; Title: DNA polymerase I gene of *Saccharomyces cerevisiae*: nucleotide sequence, mapping  
A; Reference number: A30211; MUID:88234507; PMID:3287376  
A; Accession: A30211  
A; Molecule type: DNA  
A; Residues: 1-758, 'IV', 761-1468 <PIZ>  
A; Cross-references: EMBL:J03268; NID:G172208; PIDN:AAA34888.1; PID:G172209  
R; Note: the authors translated the codon CAT for residue 653 as Ile  
R; Thomas, D.; Jacquemin, I.; Surdin-Kerjan, Y.  
Mol. Cell. Biol. 12, 1719-1727, 1992  
A; Title: MET4, a leucine zipper protein, and centromere-binding factor 1 are both required  
A; Reference number: A42246; MUID:92195318; PMID:1549123  
A; Accession: C42246  
A; Molecule type: DNA  
A; Residues: 1-6 <THO>  
A; Cross-references: EMBL:Z12126  
A; Note: sequence extracted from NCBI backbone (NCBIN:89798, NCBIP:89808)  
R; Mountain, H.A.; Byström, A.S.; Korch, C.  
Mol. Microbiol. 7, 215-228, 1993  
A; Title: The general amino acid control regulates MET4, which encodes a methionine-pathway  
A; Reference number: S31199; MUID:93188702; PMID:8446029  
A; Accession: S31201  
A; Molecule type: DNA  
A; Residues: 1-26 <MOU>  
A; Cross-references: EMBL:Z12126; NID:G3929; PIDN:CAA78111.1; PID:G3932  
R; Saiz, J.E.; Buित्रago, M.J.; Soler-Mira, A.; del Rey, F.; Revuelta, J.L.  
submitted to the Protein Sequence Database, April 1996  
A; Reference number: S63037  
A; Accession: S63042  
A; Molecule type: DNA  
A; Residues: 1-1468 <SAW>  
A; Cross-references: EMBL:Z71378; NID:G1302015; PIDN:CAA95978.1; PID:G1302016; MIPS:YNL10  
A; Experimental source: strain S288C  
R; Saiz, J.E.; Buित्रago, M.J.; Soler-Mira, A.; del Rey, F.; Revuelta, J.L.  
Yeast 12, 403-409, 1996  
A; Title: The sequence of a 21.3 kb DNA fragment from the left arm of yeast chromosome XI  
A; Reference number: S63948; MUID:96267765; PMID:8701612  
A; Accession: S63952  
A; Status: nucleic acid sequence not shown; translation not shown  
A; Molecule type: DNA  
A; Residues: 1-1468 <SAF>  
A; Cross-references: EMBL:Z50161; NID:G929846; PIDN:CAA90524.1; PID:G929851  
R; Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1995  
C; Genetics:  
A; Gene: SGD:POL1; CDC17; CRT5; HPR3  
A; Cross-references: SGD:S0005046; MIPS:YNL102w  
A; Map position: 14L  
C; Superfamily: DNA polymerase  
C; Keywords: nucleotidyltransferase

Query Match 26.7%; Score 55.5; DB 2; Length 1468;  
Best Local Similarity 29.4%; Pred. No. 26;  
Matches 15; Conservative 4; Mismatches 21; Indels 11; Gaps 1;

QY 2 LSNRLILL-----NKAALRKAHTSMVRNFRYKPVQSOLKPRDLC 41  
DB 533 MENLRCLSLISQTLMNPKENKQEIIVITLSAYRISLDSPIPENIKPDLLC 583

RESULT 9  
F81900  
biochemical protein NMA1316 [imported] - Neisseria meningitidis (strain 22491 serogroup

A:Genetics: BH2178  
A:Gene: BH2178  
Query Match 25.7%; Score 53.5; DB 2; Length 383;  
Best Local Similarity 42.9%; Pred. No. 11;  
Matches 15; Conservative 3; Mismatches 14; Indels 3; Gaps 1;  
QY 2 LSNRLILNKALRKHAHTSMVRNFRYKGPVQSOLK 36  
DB 114 IQNKILLNSAATEKATGSPARG---GKPTTATK 145  
RESULT 12  
S31290  
Cyclin B5 - yeast (Saccharomyces cerevisiae)  
N:Alternate names: protein P9642.8; protein YPR120c  
C:Species: Saccharomyces cerevisiae  
C:Date: 28-May-1993 #sequence\_revision 28-May-1993 #text\_change 12-Jul-2004  
C:Accession: S31290; S36228; S69013  
R:Epstein, C.B.; Cross, F.R.  
Genes Dev. 6, 1695-1706, 1992  
A:Title: CLB5: a novel B cyclin from budding yeast with a role in S phase.  
A:Reference number: S31290; MUID:92387544; PMID:1387626  
A:Accession: S31290  
A:Molecule type: DNA  
A:Residues: 1-435 <EPS>  
A:Cross-references: UNIPROT:P30283; EMBL:M91209; NID:g171238; PIDN:AAA34503.1; PID:g1712  
R:Schweb, E.; Nasmyth, K.  
Genes Dev. 7, 1160-1175, 1993  
A:Title: CLB5 and CLB6, a new pair of B cyclins involved in DNA replication in Saccharom  
A:Reference number: S36228; MUID:93307652; PMID:8319908  
A:Accession: S36228  
A>Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-435 <SCH>  
A:Cross-references: EMBL:X70435; NID:g396496; PIDN:CAA49893.1; PID:g396497  
R:Johnson, D.  
submitted to the EMBL Data Library, November 1995  
A:Description: The sequence of S. cerevisiae cosmid 9642.  
A:Reference number: S69014  
A:Accession: S69013  
A:Molecule type: DNA  
A:Residues: 1-435 <JOH>  
A:Cross-references: EMBL:U40828; NID:g1066467; PIDN:AAB68061.1; PID:g1066475; MIPS:YPR12  
C:Genetics:  
A:Gene: SGD:CLB5; CLB5  
A:Cross-references: SGD:S0006324  
A:Map position: 16R  
C:Superfamily: cyclin, A/B/D/E type  
C:Keywords: cell cycle control  
Query Match 25.5%; Score 53; DB 2; Length 435;  
Best Local Similarity 46.4%; Pred. No. 15;  
Matches 13; Conservative 4; Mismatches 11; Indels 0; Gaps 0;  
QY 3 SNRLILNKALRKHAHTSMVRNFRYKGP 30  
DB 28 SNLKILQNKRLSKNDSSSKQVQDSKP 55  
RESULT 13  
B81708  
translation elongation factor P TC0398 [imported] - Chlamydia muridarum (strain Nigg)  
C:Species: Chlamydia muridarum, Chlamydia trachomatis MoPn  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004  
C:Accession: B81708  
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,  
C.; Dodson, R.; Gwin, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,  
Nucleic Acids Res. 28, 1397-1406, 2000  
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.  
A:Reference number: A81500; MUID:20150255; PMID:10684935  
A:Accession: B81708  
A>Status: preliminary  
A:Molecule type: DNA

A:Molecule type: DNA  
A:Residues: 1-185 <TET>  
A:Cross-references: UNIPROT:Q9PKR6; GB:AE002306; GB:AE002160; NID:g7190428; PIDN:AAF3925;  
A:Experimental source: strain Nigg (MoPn)  
C:Genetics:  
A:Gene: TC0398  
C:Superfamily: translation elongation factor EF-P  
Query Match 25.2%; Score 52.5; DB 2; Length 185;  
Best Local Similarity 42.4%; Pred. No. 7.1;  
Matches 14; Conservative 9; Mismatches 7; Indels 3; Gaps 2;  
QY 11 KAALRKHAHTSMV--RNFRYKGPV-OSOLKPRDL 40  
DB 38 KVALQAQSDVVVERNFKAGQEVKEAQPEPRNL 70  
RESULT 14  
T52305  
En/Spm-like transposon protein [imported] - Arabidopsis thaliana  
N:Alternate names: protodermal factor 1  
C:Species: Arabidopsis thaliana (mouse-ear cross)  
C:Date: 20-Oct-2000 #sequence\_revision 20-Oct-2000 #text\_change 09-Jul-2004  
C:Accession: T52305; H84858  
R:Abe, M.; Takahashi, T.; Komeda, Y.  
Plant Cell Physiol. 40, 571-580, 1999  
A:Title: Cloning and characterization of an L1 layer-specific gene in Arabidopsis thaliana  
A:Reference number: Z26029; MUID:99412619; PMID:10483121  
A:Accession: T52305  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-306 <ABE>  
A:Cross-references: UNIPROT:Q9S728; EMBL:AF141375; PIDN:AAD33868.1  
A:Experimental source: cultivar Columbia; meristem L1 layer  
R:Liu, X.; Kaul, S.; Rounsley, S.D.; Shea, T.B.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;  
euss, D.; Niernman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487; PMID:10617197  
A:Accession: H84858  
A:Molecule type: DNA  
A:Residues: 1-306 <STO>  
A:Cross-references: GB:AE002093; NID:g4512671; PIDN:AAD21725.1; GSPDB:GN00139  
C:Genetics:  
A:Gene: PDF1; At2g42840  
A:Map position: 2  
Query Match 24.8%; Score 51.5; DB 2; Length 306;  
Best Local Similarity 41.4%; Pred. No. 17;  
Matches 12; Conservative 5; Mismatches 5; Indels 7; Gaps 1;  
QY 10 NKAALRKHAHTSMVRNFRYKGPVQSOLKPR 38  
DB 284 NKAATKQAHTFKLAN-----EGRLKPR 305  
RESULT 15  
F75504  
probable pilin, type IV - Deinococcus radiodurans (strain R1)  
C:Species: Deinococcus radiodurans  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C:Accession: F75504  
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
M.; Shen, M.; Vanathavan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A:Reference number: A75250; MUID:20036896; PMID:10567266  
A:Accession: F75504  
A>Status: preliminary  
A:Molecule type: DNA



A:Gene: SPAC29E6.01; SPDB:SPAC30.05  
A:Map position: 1  
A:Introns: 43/1; 74/3  
C:Superfamily: WD repeat homology

Query Match 24.5%; Score 51; DB 2; Length 506;  
Best Local Similarity 44.4%; Pred. No. 35;  
Matches 12; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 3 SNLRLLNKAALRKHAHTSMVRNFRYCK 29  
DB 172 SNGRPLLNWSLYLKEHAHLDSNRHGR 198

RESULT 19  
T23130  
hypothetical protein T28B8.4 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T23130; T25403  
R:White, S.  
submitted to the EMBL Data Library, December 1997  
A:Reference number: Z19690  
A:Accession: T23130  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1779 <WIL>  
A:Cross-references: UNIPROT:O18150; EMBL:AL021066; PIDN:CAA15925.1; GSPDB:GN00019; CESP:  
A:Experimental source: clone H31B20  
R:White, S.  
submitted to the EMBL Data Library, October 1996  
A:Reference number: Z20029  
A:Accession: T25403  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1779 <W12>  
A:Cross-references: EMBL:Z81133; PIDN:CAB03445.1; GSPDB:GN00019; CESP:T28B8.4  
A:Experimental source: clone T28B8  
C:Genetics:  
A:Gene: CESP:T28B8.4  
A:Map position: 1  
A:Introns: 161/2; 223/2; 309/3; 332/2; 547/3; 603/1; 657/3; 745/2; 802/1; 856/1; 1031/1;  
C:Superfamily: Caenorhabditis elegans hypothetical protein T28B8.3

Query Match 24.5%; Score 51; DB 2; Length 1779;  
Best Local Similarity 46.7%; Pred. No. 1.4e+02;  
Matches 14; Conservative 4; Mismatches 8; Indels 4; Gaps 1;

QY 1 MSLNRLIL----LNKAALRKHAHTSMVRNFR 26  
DB 1718 ILKPLRLSTSSNKAARKQTMTVTRPR 1747

RESULT 20  
T32637  
hypothetical protein F42G8.10 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T32637  
R:Gatung, S.; Holmes, A.  
submitted to the EMBL Data Library, December 1997  
A:Description: The sequence of C. elegans cosmid F42G8.  
A:Reference number: Z21203  
A:Accession: T32637  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-181 <GAT>  
A:Cross-references: UNIPROT:O44509; EMBL:AF038618; PIDN:AAB92069.1; GSPDB:GN00022; CESP:  
A:Experimental source: strain Bristol N2; clone F42G8  
C:Genetics:  
A:Gene: CESP:F42G8.10  
A:Map position: 4  
A:Introns: 41/2; 81/3; 128/2; 161/2

C:Superfamily: Caenorhabditis elegans hypothetical protein F42G8.10

Query Match 23.8%; Score 49.5; DB 2; Length 181;  
Best Local Similarity 42.9%; Pred. No. 19;  
Matches 12; Conservative 4; Mismatches 11; Indels 1; Gaps 1;

QY 13 ALRKAHTSMVRNFRYCKP-VQSOLKPRD 39  
DB 135 ALRKAHTSMVRNFRYCKP-VQSOLKPRD 162

RESULT 21  
I40646  
sensor-like protein - Coxiella burnetii  
C:Species: Coxiella burnetii  
C:Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 09-Jul-2004  
C:Accession: I40646  
R:Mo, Y.Y.; Mallavia, L.P.  
Gene 151, 185-190, 1994  
A:Title: A coxiella burnetii gene encodes a sensor-like protein.  
A:Reference number: I40646; MUID:95129857; PMID:7828872  
A:Accession: I40646  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-425 <RES>  
A:Cross-references: UNIPROT:Q45965; EMBL:U07186; NID:g460628; PIDN:AAA81939.1; PID:g46062;  
C:Genetics:  
A:Gene: grsa  
C:Superfamily: envz protein; sensor histidine kinase homology  
C:Keywords: autophosphorylation; phosphohistidine; phosphoprotein  
F196-421/Domain: sensor histidine kinase homology <SHK>  
F1227/Binding site: phosphate (His) (covalent) (by autophosphorylation) #status predicted

Query Match 23.8%; Score 49.5; DB 2; Length 425;  
Best Local Similarity 34.3%; Pred. No. 47;  
Matches 12; Conservative 11; Mismatches 11; Indels 1; Gaps 1;

QY 5 LRILLNKAALRKHAHTSMVRN-FRYCKPVQSOLKPR 38  
DB 311 IKILGRASALKRAFTNLNNAIRYAKNVNRIQWR 345

RESULT 22  
H86026  
hypothetical protein Z4949 [imported] - Escherichia coli (strain O157:H7, substrain EDL93;  
C:Species: Escherichia coli  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004  
C:Accession: H86026  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,  
illier, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: A85480; MUID:21074935; PMID:11206551  
A:Accession: H86026  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-178 <STO>  
A:Cross-references: UNIPROT:Q8X5L6; GB:AE005174; NID:g12518253; PIDN:AAG58676.1; GSPDB:G  
A:Experimental source: strain O157:H7, substrain EDL933  
C:Genetics:  
A:Gene: Z4949

Query Match 23.6%; Score 49; DB 2; Length 178;  
Best Local Similarity 33.3%; Pred. No. 21;  
Matches 11; Conservative 6; Mismatches 16; Indels 0; Gaps 0;

QY 3 SNLRLLNKAALRKHAHTSMVRNFRYCKPVQSOL 35  
DB 79 ANCHIRLHQOALPDGAHILINNFRIGSQVQDDI 111

RESULT 23  
F91180

hypothetical protein EC6414 [imported] - Escherichia coli (strain O157:H7, substrain R1)  
C;Species: Escherichia coli  
C;Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 09-Jul-2004  
C;Accession: F91180  
R;Hayaashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;  
Gawwara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen  
A;Reference number: A99629; MUID:21156231; PMID:11258796  
A;Accession: F91180  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-178 <HAY>  
A;Cross-references: UNIPROT:Q8X5L6; GB:BA000007; PIDN:BA337837.1; PID:g13363888; GSPDB:G  
A;Experimental source: strain O157:H7, substrain R1MD 050952  
C;Genetics:  
A;Gene: EC6414

Query Match 23.6%; Score 49; DB 2; Length 178;  
Best Local Similarity 33.3%; Pred. No. 21;  
Matches 11; Conservative 6; Mismatches 16; Indels 0; Gaps 0;

QY 3 SNLRILLKALRKAHTSMVRNFRYKPVQSQL 35  
Db 79 ANCHIRLHOQALPDGAHILNNFRIGSQVQDDI 111

RESULT 24  
AD3351  
spou family of rRNA methylases [imported] - Brucella melitensis (strain 16M)  
C;Species: Brucella melitensis  
C;Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 09-Jul-2004  
C;Accession: AD3351  
R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,  
.; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letes  
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitens  
A;Reference number: AD3252; PMID:11756688  
A;Accession: AD3351  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-276 <KUR>  
A;Cross-references: UNIPROT:Q8YHK3; UNIPROT:Q8G0A5; GB:AE008917; PIDN:AAL51975.1; PID:g1  
A;Experimental source: strain 16M  
C;Genetics:  
A;Gene: BMEI0794  
A;Map position: 1  
C;Superfamily: conserved hypothetical protein MTH1849

Query Match 23.6%; Score 49; DB 2; Length 276;  
Best Local Similarity 31.6%; Pred. No. 35;  
Matches 12; Conservative 7; Mismatches 19; Indels 0; Gaps 0;

QY 1 MSLNRLILKALRKAHTSMVRNFRYKPVQSQLKPR 38  
Db 220 MNNLRVAVLTRAGFAGAEKLRGVVTSLDVFTPKPR 257

RESULT 25  
T32702  
hypothetical protein C14C6.13 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T32702  
R;David, M.; Wohldmann, P.; Bauer, C.; Antoniou, B.  
submitted to the EMBL Data Library, December 1997  
A;Description: The sequence of C. elegans cosmid C14C6.  
A;Reference number: Z21210  
A;Accession: T32702  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-300 <DAV>  
A;Cross-references: UNIPROT:O44679; EMBL:AF039051; PIDN:AAB94269.1; GSPDB:GN00023; CESP:

A;Experimental source: strain Bristol N2; clone C14C6  
C;Genetics:  
A;Gene: CESP:C14C6.13  
A;Map position: 5  
A;Introns: 92/1; 129/2; 145/3; 267/3  
C;Superfamily: Caenorhabditis elegans hypothetical protein C14C6.13

Query Match 23.6%; Score 49; DB 2; Length 300;  
Best Local Similarity 28.6%; Pred. No. 38;  
Matches 10; Conservative 11; Mismatches 14; Indels 0; Gaps 0;

QY 2 LSNRLILKALRKAHTSMVRNFRYKPVQSQLK 36  
Db 199 LDMRELQSQALQKCGSETRDPTNGFYRPINSEIQ 233

RESULT 26  
AB0194  
pseudouridylate synthase [EC 4.2.1.70] [imported] - Yersinia pestis (strain CO92)  
C;Species: Yersinia pestis  
C;Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Jul-2004  
C;Accession: AB0194  
R;Farkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;  
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; I  
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, I  
Nature 413, 523-527, 2001  
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A;Reference number: AB0001; MUID:21470413; PMID:11586360  
A;Accession: AB0194  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-320 <KUR>  
A;Cross-references: UNIPROT:Q8ZFUL; GB:AL590842; PIDN:CAC90413.1; PID:g15979629; GSPDB:G  
C;Genetics:  
A;Gene: YP01591  
C;Superfamily: conserved hypothetical protein HI0176  
C;Keywords: carbon-oxygen lyase; hydro-lyase

Query Match 23.6%; Score 49; DB 2; Length 320;  
Best Local Similarity 35.5%; Pred. No. 41;  
Matches 11; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY 6 RILLKALRKAHTSMVRNFRYKPVQSQLK 36  
Db 277 RLFLHAAALRFEHPNTGETMRTIEAPLDNQLR 307

RESULT 27  
T25678  
hypothetical protein F08D12.10 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T25678  
R;Je, T.; Waterston, R.  
submitted to the EMBL Data Library, December 1996  
A;Description: The sequence of C. elegans cosmid F08D12.  
A;Reference number: Z20068  
A;Accession: T25678  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-353 <LET>  
A;Cross-references: UNIPROT:P91233; EMBL:U80840; PIDN:AAB37930.1; GSPDB:GN00020; CESP:FO  
A;Experimental source: strain Bristol N2; clone F08D12  
C;Genetics:  
A;Gene: CESP:F08D12.10  
A;Map position: 2  
A;Introns: 29/2

Query Match 23.6%; Score 49; DB 2; Length 353;  
Best Local Similarity 28.6%; Pred. No. 45;  
Matches 16; Conservative 10; Mismatches 10; Indels 20; Gaps 3;

QY 4 NLRILINK-----AALRKAHTSMV-RNFRY-----GKPVQSQLKPRD 39



A:Residues: 1-60,'A' <LIS>  
R:Horowitz, H.; Christie, G.E.; Platt, T.  
J. Mol. Biol. 156, 245-256, 1982  
A:Title: Nucleotide sequence of the trpD gene, encoding anthranilate synthetase component II  
A:Reference number: 156374; MUID:82216842; PMID:6283099  
A:Accession: 156374  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-191,'H',193-531 <RES>  
A:Cross-references: EMBL:V00367; NID:G43193; PIDN:CAA23665.1; PID:G43194  
C:Genetics:  
A:Gene: trpG-trpD  
A:Map position: 28 min  
C:Function: heterotrimer; two components I and two components II  
C:Function: <COM>  
A:Description: Anthranilate synthase catalyzes biosynthesis of anthranilate from chorismate  
A:Pathway: tryptophan biosynthesis  
A:Note: Component II is larger and its carboxyl-terminal two-thirds has anthranilate phosphatase activity  
C:Function: <COI>  
A:Description: Component I catalyzes the formation of anthranilate using ammonia rather than ATP  
C:Function: <COII>  
A:Description: Component II provides the glutamine amidotransferase activity  
C:Superfamily: trpG-trpD bifunctional enzyme; trpD homology; trpG homology  
C:Keywords: carbon-carbon lyase; glycosyltransferase; oxo-acid-lyase; pentosyltransferase  
F:4-187/Domain: trpG homology <TRG>  
F:203-530/Domain: trpD homology <TRD>  
F:84/Active site: Cys #status predicted

Query Match 23.6%; Score 49; DB 1; Length 531;  
Best Local Similarity 48.0%; Pred. No. 71;  
Matches 12; Conservative 2; Mismatches 7; Indels 4; Gaps 1;

Qy 16 KAHTSMVRNFRYGRKPVQSQLKPRDL 40  
Db 352 KYHT----GFRHAMPVRQQLKTRTL 372

RESULT 30  
AE0653  
anthranilate synthase component II, anthranilate phosphoribosyltransferase [imported] - Escherichia coli  
C:Species: Salmonella enterica subsp. enterica serovar Typhi  
A:Note: This species has also been called Salmonella typhi  
C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
C:Accession: AE0653  
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; J. Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.  
Nature 413, 848-852, 2001  
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;  
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar  
A:Reference number: AB0502; MUID:21534947; PMID:11677608  
A:Accession: AE0653  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-531 <PAR>  
A:Cross-references: GB:AL513382; PIDN:CAD08408.1; PID:gi6502451; GSPDB:GN00176  
C:Genetics:  
A:Gene: STX1327  
C:Superfamily: trpG-trpD bifunctional enzyme; trpD homology; trpG homology

Query Match 23.6%; Score 49; DB 2; Length 531;  
Best Local Similarity 48.0%; Pred. No. 71;  
Matches 12; Conservative 2; Mismatches 7; Indels 4; Gaps 1;

Qy 16 KAHTSMVRNFRYGRKPVQSQLKPRDL 40  
Db 352 KYHT----GFRHAMPVRQQLKTRTL 372

RESULT 31  
E85761  
hypochemical protein trpD [imported] - Escherichia coli (strain O157:H7, substrain EDL93)  
C:Species: Escherichia coli

C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004  
C;Accession: E85761  
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glaesner, J.D.; Rose, D.J.; Mayhew  
Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamoumis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A;Reference number: A85480; MUID:21074935; PMID:11206551  
C;Accession: E85761  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-531 <STO>  
A;Cross-references: UNIPROT:Q8X7B8; GB:AE005174; NID:gl2515535; PIDN:AAG56553.1; GSPDB:G  
A;Experimental source: strain O157:H7, substrain EDL933  
C;Genetics:  
C;Superfamily: trpG-trpD bifunctional enzyme; trpD homology; trpG homology  
Query Match 23.6%; Score 49; DB 2; Length 531;  
Best Local Similarity 48.0%; Pred. No. 71;  
Matches 12; Conservative 2; Mismatches 7; Indels 4; Gaps 1;  
QY 16 KAHTSMVNRNFRYKPKVQSOLKPRDL 40  
DB 352 KYHT----GFRHAMPVRQQLKTRTL 372  
RESULT 32  
C90858  
anthranilate synthase component II [imported] - Escherichia coli (strain O157:H7, substr  
C;Species: Escherichia coli  
C;Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 09-Jul-2004  
C;Accession: C90858  
R;Hayashi, T.; Makino, K.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
Sasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gend  
A;Reference number: A99629; MUID:21156231; PMID:11258796  
C;Accession: C90858  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-531 <HAY>  
A;Cross-references: UNIPROT:Q8X7B8; GB:BA000007; PIDN:BA035258.1; PID:gl3361300; GSPDB:G  
A;Experimental source: strain O157:H7, substrain RIMD 0509952  
C;Genetics:  
C;Superfamily: trpG-trpD bifunctional enzyme; trpD homology; trpG homology  
Query Match 23.6%; Score 49; DB 2; Length 531;  
Best Local Similarity 48.0%; Pred. No. 71;  
Matches 12; Conservative 2; Mismatches 7; Indels 4; Gaps 1;  
QY 16 KAHTSMVNRNFRYKPKVQSOLKPRDL 40  
DB 352 KYHT----GFRHAMPVRQQLKTRTL 372  
RESULT 33  
F83354  
probable sulfatase PA2333 [imported] - Pseudomonas aeruginosa (strain PA01)  
C;Species: Pseudomonas aeruginosa  
C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004  
C;Accession: F83354  
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, K.R.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
A;Reference number: A82950; MUID:20437337; PMID:10984043  
C;Accession: F83354  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-538 <STO>  
A;Cross-references: UNIPROT:Q911E5; GB:AE004659; GB:AE004091; NID:g9948361; PIDN:AAG0572

A;Experimental source: strain PA01

C;Genetics:

A;Gene: PA2333

Query Match 23.6%; Score 49; DB 2; Length 538;

Best Local Similarity 27.1%; Pred. No. 72;

Matches 13; Conservative 6; Mismatches 7; Indels 22; Gaps 1;

QY 12 AALRKQHTSMV-----RNFYKPKVQSOLK 37

DB 487 AAVREAHAGMLFDWLRLKRRRTTISNAEIDLRCQAFRYGPEGRLLVP 534

RESULT 34

A29646

invasin - Yersinia pseudotuberculosis

C;Species: Yersinia pseudotuberculosis

C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004

C;Accession: A29646; S12543

R;Isberg, R.R.; Voorhis, D.L.; Falkow, S.

Cell 50, 769-778, 1987

A;Title: Identification of invasin: a protein that allows enteric bacteria to penetrate

A;Reference number: A29646; MUID:87301720; PMID:3304658

A;Accession: A29646

A;Molecule type: DNA

A;Residues: 1-986 <ISB>

A;Cross-references: UNIPROT:P11922; GB:M17448; NID:gl155439; PIDN:AAA27632.1; PID:gl155440

R;Leong, J.M.; Fournier, R.S.; Isberg, R.R.

EMBO J. 9, 1979-1989, 1990

A;Title: Identification of the integrin binding domain of the Yersinia pseudotuberculosis

A;Reference number: S12543; MUID:90262235; PMID:1693333

A;Accession: S12543

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 795-986 <LEO>

C;Superfamily: Invasin

Query Match 23.6%; Score 49; DB 1; Length 986;

Best Local Similarity 31.0%; Pred. No. 1.4e+02;

Matches 9; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

QY 9 LNKAAURKAHTSMVRNFRYKPKVQSOLK 37

DB 335 MGKSKKHETQWNLQMYRLGSGFSQSLSP 363

RESULT 35

AI0802

Div protein [imported] - Salmonella enterica subsp. enterica serovar Typhi (strain CT18)

C;Species: Salmonella enterica subsp. enterica serovar Typhi

A;Note: this species has also been called Salmonella typhi

C;Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002

C;Accession: AI0802

R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,

th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,

S.; Moule, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;

A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov

A;Reference number: AB0502; MUID:21534947; PMID:11677608

C;Accession: AI0802

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-333 <PAR>

A;Cross-references: GB:AL513382; PIDN:CAD07603.1; PID:gl6503595; GSPDB:GN00176

C;Genetics:

A;Gene: STY2602

Query Match 23.3%; Score 48.5; DB 2; Length 333;

Best Local Similarity 37.5%; Pred. No. 50;

Matches 15; Conservative 6; Mismatches 18; Indels 1; Gaps 1;

QY 1 MLNSRLILINKAALRKAHTSMVRNFRYKPKVQSOLKPRDL 40

C;Accession: A71033



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 19, 2004, 02:22:26 ; Search time 123.488 Seconds  
(without alignments)  
191.033 Million cell updates/sec

Title: US-08-765-244-1  
Perfect score: 208  
Sequence: 1 MSLNRLILLNKAALRKAHTS.....VNFYRGKPVQSLKPRDLC 41

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Uniprot 02: \*  
1: uniprot\_sprot: \*  
2: uniprot\_trembl: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	179	86.1	354	1	OTC RAT	P00481 rattus norv
2	144	69.2	351	2	Q8RIA8	Q8RIA8 mus musculu
3	144	69.2	354	1	OTC MOUSE	P11725 mus musculu
4	144	69.2	354	2	BAC34465	BAC34465 mus muscu
5	125	60.1	354	1	OTC HUMAN	P00480 homo sapien
6	114	54.8	354	2	Q9N1U7	Q9N1U7 bos taurus
7	112	53.8	41	2	Q63786	Q63786 rattus sp.
8	92	44.2	354	2	Q9IAU8	Q9IAU8 trachemys s
9	57	27.4	350	1	OTC RANCA	P11326 rana catesb
10	57	27.4	351	2	Q97W55	Q97W55 sulfolobus
11	56	26.9	356	2	Q9IAV0	Q9IAV0 sceloporus
12	56	26.9	1286	2	Q23351	Q23351 arabidopsis
13	55.5	26.7	302	2	Q89PI4	Q89PI4 bradyrhizob
14	55.5	26.7	795	2	Q74H89	Q74H89 geobacter s
15	55.5	26.7	795	2	AAR33338	AAR33338 geobacter
16	55.5	26.7	1468	1	DPOA YEAST	P13382 saccharomyc
17	55	26.4	258	2	Q6P0K2	Q6P0K2 brachydanio
18	55	26.4	258	2	AAH65586	AAH65586 brachydan
19	55	26.4	338	2	Q6CJVS	Q6CJVS kluyveromyc
20	54.5	26.2	138	2	Q9JUG9	Q9JUG9 neisseria m
21	54	26.0	236	1	R92 LOTJA	Q9JBS6 lotus japon
22	54	26.0	271	2	Q732D6	Q732D6 mycobacteri
23	54	26.0	271	2	AAS03984	AAS03984 mycobacte
24	54	26.0	457	2	Q8AVS9	Q8AVS9 xenopus lae
25	54	26.0	544	1	A37C_DROLE	Q96570 drosophila
26	54	26.0	769	2	Q52393	Q52393 pseudomonas
27	54	26.0	846	2	Q6GGQ76	Q6GGQ76 xenopus lae
28	54	26.0	1147	2	Q7Q642	Q7Q642 anopheles g
29	53.5	25.7	260	2	O50746	O50746 borrelia bu
30	53.5	25.7	333	2	Q9KAV8	Q9KAV8 bacillus ha
31	53.5	25.7	1113	2	Q7SGU4	Q7SGU4 neurospora

RESULT 1		ALIGNMENTS	
OTC_RAT	OTC_RAT	STANDARD;	PRT; 354 AA.
AC P00481; Q63407;			
DT 21-JUL-1986 (Rel. 01, Created)			
DT 21-JUL-1986 (Rel. 01, Last sequence update)			
DT 05-JUL-2004 (Rel. 44, Last annotation update)			
DE Ornithine carbamoyltransferase, mitochondrial precursor (EC 2.1.3.3)			
DE (OTCase) (Ornithine transcarbamylase).			
GN Name=Otc;			
OS Rattus norvegicus (Rat).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX NCBI_TaxID=10116;			
RN [1]			
RP SEQUENCE FROM N.A.			
RX MEDLINE=85063800; PubMed=6095294;			
RA Takiguchi M., Miura S., Mori M., Tatibana M., Nagata S., Kaziro Y.;			
RT "Molecular cloning and nucleotide sequence of cDNA for rat ornithine			
RT carbamoyltransferase precursor.";			
RL Proc. Natl. Acad. Sci. U.S.A. 81:7412-7416(1984).			
RN [2]			
RP SEQUENCE FROM N.A.			
RC STRAIN=Wistar; TISSUE=Liver;			
RX MEDLINE=87317609; PubMed=3476935;			
RA Takiguchi M., Murakami T., Miura S., Mori M.;			
RT "Structure of the rat ornithine carbamoyltransferase gene, a large, X			
RT chromosome-linked gene with an atypical promoter.";			
RL Proc. Natl. Acad. Sci. U.S.A. 84:6136-6140(1987).			
RN [3]			
RP SEQUENCE FROM N.A.			
RX MEDLINE=85215524; PubMed=3839075;			
RA Kraus J.P., Hodges P.E., Williamson C.L., Horwich A.L., Kalousek F.,			
RA Williams K.R., Rosenberg L.E.;			
RT "A cDNA clone for the precursor of rat mitochondrial ornithine			
RT transcarbamylase: comparison of rat and human leader sequences and			
RT conservation of catalytic sites.";			
RL Nucleic Acids Res. 13:943-952(1985).			
RN [4]			
RP SEQUENCE FROM N.A.			
RX MEDLINE=85203360; PubMed=3838931;			
RA McIntyre P., Graf L., Mercer J.F.B., Wake S.A., Hudson P.J.,			
RA Hoogenraad N.;			
RT "The primary structure of the imported mitochondrial protein,			
RT ornithine transcarbamylase from rat liver: mRNA levels during			
RT ontogeny.";			
RL DNA 4:147-156(1985).			
RN [5]			
RP SEQUENCE OF 1-102 FROM N.A.			
RX MEDLINE=85051832; PubMed=6548714;			
RA McIntyre P., Graf L., Mercer J., Peterson G., Hudson P.J.,			
RA Hoogenraad N.;			
RT "A highly basic N-terminal extension of the mitochondrial matrix			
RT enzyme ornithine transcarbamylase from rat liver.";			

32	53	25.5	236	2	Q6LG10	Q6lg10 photobacter
33	53	25.5	236	2	CAG23600	Cag23600 photobact
34	53	25.5	355	2	Q6IQS0	Q6iqs0 brachydanio
35	53	25.5	355	2	Q6PBM0	Q6pbm0 brachydanio
36	53	25.5	355	2	AAH59660	Aah59660 brachydan
37	53	25.5	435	1	CGS5 YEAST	P30283 saccharomyc
38	53	25.5	457	2	Q7VV31	Q7vv31 bordetella
39	53	25.5	597	1	IF2P METAC	Q8tql5 methanosarc
40	52.5	25.2	140	2	Q99A14	Q99a14 bovine vira
41	52.5	25.2	141	2	Q8DG64	Q8dgl5 synechococc
42	52.5	25.2	162	2	Q6NFT2	Q6nit2 corynebacte
43	52.5	25.2	162	2	CAE50334	Caes0334 corynebac
44	52.5	25.2	185	1	EPF1 CHLMU	Q9pkre chlamydia m
45	52.5	25.2	475	2	Q9NGB5	Q9ngb5 drosophila

```

RL FEBS Lett. 177:41-46(1984).
RN [6]
RP SEQUENCE OF 33-56; 293-302; 307-317 AND 321-329.
RC TISSUE=Liver;
RX MEDLINE=88268748; PubMed=3390141;
RA Aoki Y., Sunaga H., Suzuki K.T.;
RT "A cadmium-binding protein in rat liver identified as ornithine
RL Biochem. J. 250:735-742(1988).
CC -1- CATALYTIC ACTIVITY: Carbamoyl phosphate + L-ornithine = phosphate
CC + L-citrulline.
CC -1- PATHWAY: Second step in urea cycle, arginine biosynthesis.
CC -1- SUBUNIT: Homotrimer.
CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -1- SIMILARITY: Belongs to the ATCase/OTCase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K03040; AAA41768.1; -.
DR EMBL; M16933; AAA41769.1; -.
DR EMBL; M16924; AAA41769.1; JOINED.
DR EMBL; M16925; AAA41769.1; JOINED.
DR EMBL; M16926; AAA41769.1; JOINED.
DR EMBL; M16928; AAA41769.1; JOINED.
DR EMBL; M16929; AAA41769.1; JOINED.
DR EMBL; M16930; AAA41769.1; JOINED.
DR EMBL; M16932; AAA41769.1; JOINED.
DR EMBL; X01976; CAA26007.1; -.
DR EMBL; K00001; AAA41772.1; -.
DR EMBL; M1266; AAA41767.1; -.
DR EMBL; X01178; CAA25618.1; -.
DR PIR; A00563; OWET.
DR HSP; P00480; 1OTH.
DR RGD; 3236; Otc.
DR InterPro; IPR006131; Asp/Orn-bind.
DR InterPro; IPR005130; Asp/Orn_Cotranf.
DR InterPro; IPR002292; Orn_CarbAmLtrans.
DR Pfam; PF00185; OTCace; 1.
DR PRINTS; PR00100; AOTCase.
DR PRINTS; PR00102; OTCace.
DR TIGRfams; TIGR00658; orni_carb_tr; 1.
DR PROSITE; PS00097; CARBAMOYLTRANSFERASE; 1.
KW Arginine biosynthesis; Direct protein sequencing; Mitochondrion;
KW Transferase; Transit peptide; Urea cycle.
FT TRANSIT 1 32 Mitochondrion.
FT CHAIN 33 354 Ornithine carbamoyltransferase.
FT ACT_SITE 263 263 By similarity.
FT ACT_SITE 303 303 By similarity.
FT CONFLICT 39 39 G -> P (in Ref. 3; AAA41772).
FT CONFLICT 241 241 G -> S (in Ref. 3).
SQ SEQUENCE 354 AA; 39886 MW; 156B511AF7063F0C CRC64;

Query Match 86.1%; Score 179; DB 1; Length 354;
Best Local Similarity 92.9%; Pred. No. 2.8e-17;
Matches 39; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

QY 1 MLSNRLILLNKAALRKHAFTSMVRNFRYKPKVQS--QLKPRDL 40
DB 1 MLSNRLILLNKAALRKHAFTSMVRNFRYKPKVQSQVQLKGRDL 42

RESULT 2
Q8R1A8 PRELIMINARY; PRT; 351 AA.
AC Q8R1A8;

DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Otc protein.
GN Name=Otc;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Liver;
RX MEDLINE=22388257; PubMed=12477932;
RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Liver;
RA Strausberg R.;
RA Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the ATCase/OTCase family.
DR EMBL; BC024893; AAH24893.1; -.
DR HSSP; P00480; 1OTH.
DR MGD; MGI:97448; Otc.
DR GO; GO:0005739; C.mitochondrion; IDA.
DR InterPro; IPR006131; Asp/Orn-bind.
DR InterPro; IPR006130; Asp/Orn_Cotranf.
DR InterPro; IPR002292; Orn_CarbAmLtrans.
DR InterPro; IPR006132; OTCace_P.
DR Pfam; PF00185; OTCace; 1.
DR Pfam; PF02729; OTCace_N; 1.
DR PRINTS; PR00100; AOTCase.
DR PRINTS; PR00102; OTCace.
DR TIGRfams; TIGR00658; orni_carb_tr; 1.
DR PROSITE; PS00097; CARBAMOYLTRANSFERASE; UNKNOWN_1.
KW Transferase.
SQ SEQUENCE 351 AA; 39365 MW; E6B38BB2FC779F42 CRC64;

Query Match 69.2%; Score 144; DB 2; Length 351;
Best Local Similarity 78.6%; Pred. No. 3.4e-12;
Matches 33; Conservative 2; Mismatches 5; Indels 2; Gaps 1;

QY 1 MLSNRLILLNKAALRKHAFTSMVRNFRYKPKVQS--QLKPRDL 40
DB 1 MLSNRLILLNKAALRKHAFTSMVRNFRYKPKVQSQVQLKGRDL 42

RESULT 3
QTC_MOUSE STANDARD; PRT; 354 AA.
AC P11725;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Ornithine carbamoyltransferase, mitochondrial precursor (EC 2.1.3.3)

```

(OTCase) (Ornithine transcarbamylase).

DE Mus musculus (Mouse).

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; PubMed=3603027;

RX MEDLINE=87263407;

RA Veres G., Gibbs R.A., Scherer S.E., Caskey C.T.;

RT "The molecular basis of the sparse fur mouse mutation.";

RL Science 237:415-417(1987).

RN [2]

RP SEQUENCE FROM N.A.

RC MEDLINE=88157717; PubMed=2831503;

RA Scherer S.E., Veres G., Caskey C.T.;

RT "The genetic structure of mouse ornithine transcarbamylase.";

RL Nucleic Acids Res. 16:1593-1601(1988).

RN [3]

RP SEQUENCE OF 1-26 FROM N.A.

RC STRAIN=C57BL/6J;

RX MEDLINE=86224037; PubMed=3011788;

RA Veres G., Craigen W.J., Caskey C.T.;

RT "The 5' flanking region of the ornithine transcarbamylase gene contains DNA sequences regulating tissue-specific expression.";

RL J. Biol. Chem. 261:7588-7591(1986).

CC -!- CATALYTIC ACTIVITY: Carbamoyl phosphate + L-ornithine = phosphate + L-citrulline.

CC -!- PATHWAY: Second step in urea cycle, arginine biosynthesis.

CC -!- SUBUNIT: Homotrimer.

CC -!- SUBCELLULAR LOCATION: Mitochondrial matrix.

CC -!- DISEASE: Sparse fur (spf) mouse have an OTCase with an overall decrease in activity, and altered substrate affinity.

CC -!- SIMILARITY: Belongs to the ATCase/OTCase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; M17030; AAA39865.1; -.

DR EMBL; M12716; AAA39864.1; ALT\_SEQ.

DR EMBL; X07092; CAA30121.1; -.

DR EMBL; X07093; CAA30121.1; JOINED.

DR EMBL; X07094; CAA30121.1; JOINED.

DR EMBL; X07095; CAA30121.1; JOINED.

DR EMBL; X07096; CAA30121.1; JOINED.

DR EMBL; X07097; CAA30121.1; JOINED.

DR EMBL; X07098; CAA30121.1; JOINED.

DR EMBL; X07099; CAA30121.1; JOINED.

DR EMBL; X07100; CAA30121.1; JOINED.

DR F01; A43609; ORMS.

DR HSSP; P00480; 10TH.

DR SWISS-2DPAGE; P11725; MOUSE.

DR MGD; MGI:97448; Otc.

DR InterPro; IPR006131; Asp/Orn-bind.

DR InterPro; IPR006130; Asp/Orn\_COXranf.

DR InterPro; IPR002292; Orn\_carbamitrans.

DR InterPro; IPR006132; OTCase\_P.

DR Pfam; PF00185; OTCase; 1.

DR Pfam; PF02729; OTCase\_N; 1.

DR PRINTS; PR00100; AOTCASE.

DR PRINTS; PR00102; OTCASE.

DR TIGRFAMs; TIGR00658; orn\_i carb tr; 1.

DR PROSITE; PS00037; CARBAMOYLTRANSFERASE; 1.

KW Arginine biosynthesis; Mitochondrion; Transferase; Transit peptide;

KW Urea cycle.

FT TRANSIT 1 32 Mitochondrion.

CHAIN 33 354 Ornithine carbamoyltransferase.

FT ACT\_SITE 263 263 By similarity.

FT ACT\_SITE 303 303 By similarity.

FT VARIANT 117 117 H -> N (IN SPARSE FUR MOUSE).

FT CONFLICT 195 195 G -> R (in Ref. 2).

SQ SEQUENCE 354 AA; 39765 MW; 33BBE5D1E88AA196 CRC64;

Query Match 69.2%; Score 144; DB 1; Length 354;

Best Local Similarity 78.6%; Pred: No. 3.5e-12; Indels 2; Gaps 1;

Matches 33; Conservative 2; Mismatches 5;

QY 1 MLSNRLILLNKAALRKAGHTSVVRFHFWCGPVSQVQLKGRDL 40

DB 1 MLSNRLILLNKAALRKAGHTSVVRFHFWCGPVSQVQLKGRDL 42

RESULT 4

BAC34465 PRELIMINARY; PRT; 354 AA.

ID BAC34465;

AC BAC34465;

DT 14-APR-2004 (Tremblrel. 27, Created)

DT 14-APR-2004 (Tremblrel. 27, Last sequence update)

DT 14-APR-2004 (Tremblrel. 27, Last annotation update)

DE 9 days embryo whole body cDNA, RIKEN full-length enriched library.

DE clone: D030040624 product: ORNITHINE CARBAMOYLTRANSFERASE, MITOCHONDRIAL

DE (EC 2.1.3.3) (OTCase) (ORNITHINE TRANSCARBAMYLASE), full insert

DE sequence.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Whole body;

RX MEDLINE=22354683; PubMed=12466851;

RA The FANTOM Consortium.

RA The RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs.";

RL Nature 420:563-573(2002).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Whole body;

RX MEDLINE=21085660; PubMed=11217851;

RA RIKEN FANTOM Consortium;

RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690(2001).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Whole body;

RX MEDLINE=99279253; PubMed=10349636;

RA Carninci P., Hayashizaki Y.;

RT "High-efficiency full-length cDNA cloning.";

RL Meth. Enzymol. 303:19-44(1999).

RN [4]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Whole body;

RX MEDLINE=20499374; PubMed=11042159;

RA Carninci P., Shibata Y., Hayatsu M., Sugahara Y., Shibata K., Itoh M.,

RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;

RT "Normalization and subtraction of cap-trapper-selected cDNAs to

RT prepare full-length cDNA libraries for rapid discovery of new genes.";

RL Genome Res. 10:1617-1630(2000).

RN [5]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Whole body;

RX MEDLINE=20530913; PubMed=11076861;

RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,

RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,

RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,

RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,

RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,

RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,

RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;

RT "RIKEN integrated sequence analysis (RISA) system-384-format  
 RT sequencing pipeline with 384 multicapillary sequencer."  
 RL Genome Res. 10:1757-1771(2000).  
 RN [6]

RN SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Whole body;  
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,  
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,  
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,  
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Koyama M., Koya S.,  
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,  
 RA Nishii K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,  
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,  
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,  
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,  
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.,  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK050930; BAC34465.1; -  
 KW Transferase.  
 SQ SEQUENCE 354 AA; 39765 MW; 33BBE5D1E88AA196 CRC64;

Query Match 69.2%; Score 144; DB 2; Length 354;  
 Best Local Similarity 78.6%; Pred. No. 3.5e-12;  
 Matches 33; Conservative 2; Mismatches 5; Indels 2; Gaps 1;

Qy 1 MLSNLRILLKAAALRKAKTSMVRNFRYKPKVQS--QLKPRDL 40  
 Db 1 MLSNLRILLNAAALRKAGTSVVRHFWCKPKVQSVQLKGRDL 42

## RESULT 5

OTC\_HUMAN  
 ID\_ OTC\_HUMAN STANDARD; PRT; 354 AA.  
 AC P00480; G9NYJ5;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 01-OCT-2004 (Rel. 45, Last annotation update)  
 DE Ornithine carbamoyltransferase, mitochondrial precursor (EC 2.1.3.3)  
 DE (OTCase) (Ornithine transcarbamylase).  
 GN Name=OTC;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=88227905; PubMed=2836378;  
 RA Hata A., Tsuzuki T., Shimada K., Takiguchi M., Mori M., Matsuda I.;  
 RT "Structure of the human ornithine transcarbamylase gene."  
 RL J. Biochem. 103:302-308(1988).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=84196410; PubMed=6372096;  
 RA Horwich A.L., Fenton W.A., Williams K.R., Kalousek F., Kraus J.P.,  
 RA Doolittle R.F., Konigsberg W., Rosenberg L.E.;  
 RT "Structure and expression of a complementary DNA for the nuclear coded  
 RT precursor of human mitochondrial ornithine transcarbamylase."  
 RL Science 224:1068-1074(1984).  
 RN [3]  
 RP SEQUENCE OF 1-36 FROM N.A.  
 RX MEDLINE=85270440; PubMed=3895227;  
 RA Horwich A.L., Kalousek F., Rosenberg L.E.;  
 RT "Arginine in the leader peptide is required for both import and  
 RT proteolytic cleavage of a mitochondrial precursor."  
 RL Proc. Natl. Acad. Sci. U.S.A. 82:4930-4933(1985).  
 RN [4]  
 RP SEQUENCE OF 1-26 FROM N.A.  
 RX MEDLINE=87057134; PubMed=3782067;  
 RA Hata A., Tsuzuki T., Shimada K., Takiguchi M., Mori M., Matsuda I.;  
 RT "Isolation and characterization of the human ornithine

RT transcarbamylase gene: structure of the 5'-end region."  
 RL J. Biochem. 100:717-725(1986).  
 RN [5]  
 RP SEQUENCE OF 101-354 FROM N.A., AND VARIANT PHE-101.  
 RA Blechschmidt K., Nyakatura G., Menzel U., Baumgart C., Dette M.,  
 RA Jahn N., Strom T.M., Hellebrand H., Meindl A., Rosenthal A.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS).  
 RX MEDLINE=99069419; PubMed=9852088;  
 RA Shi D., Morizono H., Ha Y., Aoyagi M., Tuchman M., Allewell N.M.;  
 RT "1.85-A resolution crystal structure of human ornithine  
 RT transcarbamylase complexed with N-phosphonacetyl-L-ornithine."  
 RT Catalytic mechanism and correlation with inherited deficiency."  
 RL J. Biol. Chem. 273:34247-34254(1998).  
 RN [7]  
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).  
 RX MEDLINE=20274073; PubMed=10813810;  
 RX DOI=10.1002/(SICI)1097-0134(20000601)39:4<271::AID-PROT10-3.3.CO;2-5;  
 RA Shi D., Morizono H., Aoyagi M., Tuchman M., Allewell N.M.;  
 RT "Crystal structure of human ornithine transcarbamylase complexed with  
 RT carbamoyl phosphate and L-norvaline at 1.9 A resolution."  
 RL Proteins 39:271-277(2000).  
 RN [8]  
 RP REVIEW ON VARIANTS.  
 RX MEDLINE=93372868; PubMed=8364586;  
 RA Tuchman M.;  
 RT "Mutations and polymorphisms in the human ornithine transcarbamylase  
 RT gene."  
 RL Hum. Mutat. 2:174-178(1993).  
 RN [9]  
 RP REVIEW ON VARIANTS.  
 RX MEDLINE=95353279; PubMed=7627182;  
 RA Tuchman M., Plante R.J.;  
 RT "Mutations and polymorphisms in the human ornithine transcarbamylase  
 RT gene: mutation update addendum."  
 RL Hum. Mutat. 5:293-295(1995).  
 RN [10]  
 RP REVIEW ON VARIANTS, AND 3D-STRUCTURE MODELING.  
 RX MEDLINE=96091868; PubMed=8544185;  
 RA Tuchman M., Morizono H., Reish O., Yuan X., Allewell N.M.;  
 RT "The molecular basis of ornithine transcarbamylase deficiency:  
 RT modelling the human enzyme and the effects of mutations."  
 RL J. Med. Genet. 32:680-688(1995).  
 RN [11]  
 RP VARIANT OTC GLN-141.  
 RX MEDLINE=89008892; PubMed=3170748;  
 RA Maddalena A., Spence J.E., O'Brien W.E., Nuesbaum R.L.;  
 RT "Characterization of point mutations in the same arginine codon in  
 RT three unrelated patients with ornithine transcarbamylase deficiency."  
 RL J. Clin. Invest. 82:1353-1358(1988).  
 RN [12]  
 RP VARIANTS OTC GLN-26; PRO-45 AND GLU-216, AND VARIANT ARG-46.  
 RX MEDLINE=89345570; PubMed=2474822;  
 RA Grompe M., Muzny D.M., Caskey C.T.;  
 RT "Scanning detection of mutations in human ornithine transcarbamylase  
 RT by chemical mismatch cleavage."  
 RL Proc. Natl. Acad. Sci. U.S.A. 86:5888-5892(1989).  
 RN [13]  
 RP VARIANT OTC TRP-277.  
 RX MEDLINE=90269805; PubMed=2347583;  
 RA Finkelstein J.E., Francomano C.A., Brusilow S.W., Traystman M.D.;  
 RT "Use of denaturing gradient gel electrophoresis for detection of  
 RT mutation and prospective diagnosis in late onset ornithine  
 RT transcarbamylase deficiency."  
 RL Genomics 7:167-172(1990).  
 RN [14]  
 RP VARIANTS OTC GLN-92 AND LEU-320, AND VARIANT PRO-111.  
 RX MEDLINE=91118929; PubMed=1671317;  
 RA Grompe M., Caskey C.T., Fenwick R.G. Jr.;  
 RT "Improved molecular diagnostics for ornithine transcarbamylase  
 RT deficiency."  
 RL Am. J. Hum. Genet. 48:212-222(1991).





Q63786 PRELIMINARY; PRT; 41 AA.  
 AC Q63786;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
 DE Rat ornithine transcarbamylase gene leader sequence, partial cds.  
 DE (Fragment).  
 OS Rattus sp.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10118;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86106223; PubMed=3943133;  
 RA Kowalewski A.L., Kalousek F., Fenton W.A., Pollock R.A., Rosenberg L.E.;  
 RT "Targeting of pre-ornithine transcarbamylase to mitochondria:  
 RT Definition of critical regions and residues in the leader peptide."  
 RL Cell 44:451-459 (1986).  
 DR EMBL; M12583; AAA41770.1; -.  
 FT NON\_TER 41  
 SQ SEQUENCE 41 AA; 4934 MW; 802465955B32374B CRC64;

Query Match 53.8%; Score 112; DB 2; Length 41;  
 Best Local Similarity 61.1%; Pred. No. 1.4e-08;  
 Matches 22; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

Qy 1 MSLNRLILNKAAIRKARTSMVRNFRYKPKVQSLK 36  
 |||||  
 Db 1 MLFNRLILNNAAFRNGHNFVRNFRGQPLQNKVQ 36

RESULT 8

Q91A08 PRELIMINARY; PRT; 354 AA.  
 AC Q91A08;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Ornithine transcarbamylase.  
 GN Name=OTC;  
 OS Trachemys scripta elegans.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Testudines; Cryptodira; Testudinoidae; Emydidae; Trachemys.  
 OX NCBI\_TaxID=31138;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Shimogiri T., Koyanagi K., Mannen H., Tsuji S.;  
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: Belongs to the AtCase/OTCase family.  
 DR EMBL; AF134846; AAF61410.1; -.  
 DR HSP; P00480; 1EP9.  
 DR GO; GO:0009348; C:ornithine transcarbamylase complex; IEA.  
 DR GO; GO:0016597; F:amino acid binding; IEA.  
 DR GO; GO:0004585; F:ornithine transcarbamylase activity; IEA.  
 DR GO; GO:0016740; F:transcarbamylase activity; IEA.  
 DR GO; GO:0006520; P:amino acid metabolism; IEA.  
 DR InterPro; IPR006131; Asp/Orn-bind.  
 DR InterPro; IPR002292; Orn-carbamyltrans.  
 DR InterPro; IPR006132; OTCace\_P.  
 DR Pfam; PF00185; OTCace; 1.  
 DR Pfam; PF02729; OTCace N; 1.  
 DR PRINTS; PR00100; AOTCASE.  
 DR PRINTS; PR00102; OTCASE.  
 DR TIGRPFAMs; TIGR00658; orni carb tr; 1.  
 DR PROSITE; PS00097; CARBAMOYLTRANSFERASE; UNKNOWN\_1.  
 KW Transferrase.  
 SQ SEQUENCE 354 AA; 39958 MW; BD6A2C3AEC0F99BE CRC64;  
 Query Match 44.2%; Score 92; DB 2; Length 354;  
 Best Local Similarity 52.4%; Pred. No. 0.00013;  
 Matches 22; Conservative 6; Mismatches 12; Indels 2; Gaps 1;

Qy 1 MSLNRLILNKAAIRKARTSMVRNFRYKPKVQSLK 40  
 |||||  
 Db 1 MLFNRLILNNAATLNRSSKQLVQHFRSGQPTQTINILKGRDL 42

RESULT 9

OTC\_RANCA STANDARD; PRT; 350 AA.  
 AC P31326;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Ornithine carbamoyltransferase, mitochondrial precursor (EC 2.1.3.3)  
 DE (OTCase) (Ornithine transcarbamylase).  
 OS Rana catesbeiana (Bull frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Ranidae; Rana.  
 OX NCBI\_TaxID=8400;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=93177976; PubMed=1291156;  
 RA Helbing C., Gergely G., Atkinson B.G.;  
 RT "Sequential up-regulation of thyroid hormone beta receptor, ornithine  
 RT transcarbamylase, and carbamoyl phosphate synthetase mRNAs in the liver  
 RT of Rana catesbeiana tadpoles during spontaneous and thyroid hormone-  
 RT induced metamorphosis."  
 RL Dev. Genet. 13:289-301 (1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Iwase K., Yamauchi K., Ishikawa K.;  
 RT "Molecular cloning of bullfrog (Rana catesbeiana) ornithine  
 RT transcarbamylase and induction of its mRNA during spontaneous  
 RT metamorphosis."  
 RL Rep. Fac. Sci. Shizuoka Univ. 29:45-54 (1995).  
 CC -1- FUNCTION: OTC is necessary for the tadpoles transition from an  
 CC ammonotelic, aquatic larva to a ureotelic, terrestrial adult.  
 CC -1- CATALYTIC ACTIVITY: Carbamoyl phosphate + L-ornithine = phosphate  
 CC + L-citrulline.  
 CC -1- PATHWAY: Second step in urea cycle, arginine biosynthesis.  
 CC -1- SUBUNIT: Homotrimer.  
 CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.  
 CC -1- TISSUE SPECIFICITY: Liver.  
 CC -1- DEVELOPMENTAL STAGE: Expressed during embryonic development.  
 CC -1- INDUCTION: By thyroid hormone.  
 CC -1- SIMILARITY: Belongs to the AtCase/OTCase family.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; M95193; AAA49528.1; -.  
 DR EMBL; D38304; BAA22775.1; -.  
 DR PIR; A48421; A48421.  
 DR HSP; P00480; 1OTH.  
 DR InterPro; IPR006131; Asp/Orn-bind.  
 DR InterPro; IPR006130; Asp/Orn Cotranf.  
 DR InterPro; IPR002292; Orn-carbamyltrans.  
 DR InterPro; IPR006132; OTCace\_P.  
 DR Pfam; PF00185; OTCace; 1.  
 DR Pfam; PF02729; OTCace N; 1.  
 DR PRINTS; PR00100; AOTCASE.  
 DR PRINTS; PR00102; OTCASE.  
 DR TIGRPFAMs; TIGR00658; orni carb tr; 1.  
 DR PROSITE; PS00097; CARBAMOYLTRANSFERASE; 1.  
 KW Arginine biosynthesis; Mitochondrion; Transferase; Transit peptide;  
 Urea cycle.  
 FT TRANSIT 1 30 Mitochondrion (By similarity).



```

Query Match      26.9%; Score 56; DB 2; Length 1286;
Best Local Similarity 35.3%; Pred. No. 93;
Matches 12; Conservative 8; Mismatches 14; Indels 0; Gaps 0;

QY   3 SNLRILLNKAALRKHAHSMVRNFYRGKPVQSOLK 36
    ||||| :|: ::||| :||| |::| :|:
DB   107 SNLRRMLSSSTTKRDESLVRLNLLLVPIQDIQ 140

RESULT 13
Q89PI4          PRELIMINARY;       PRT;     302 AA.
ID Q89PI4
AC Q89PI4
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE B1r3496 protein.
GN OrderedLocusNames=b1r3496;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA110;
RX KANEKO T., NAKAMURA Y., SATO S., MINAMISAWA K., UCHIUMI T.,
RA SASAMOTO S., WATANABE A., IDEGAWA K., IRIUCHI M., KAWASHIMA K.,
RA KOHARA M., MATSUMOTO M., SHIMO S., TSURUOKA H., WADA T., YAMADA M.,
RA TABATA S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
DR EMBL; AP005947; BAC48761.1; -.
GO GO:0016020; C:membrane; IEA.
GO GO:0009306; P:protein secretion; IEA.
DR InterPro; IPRO01992; Bact_sec_systII.
DR Pfam; PF00482; GSP11_F; 1.
KW Complete proteome.
SQ SEQUENCE 302 AA; 32808 MW; A46FF9AA4B08096F2 CRC64;

Query Match      26.7%; Score 55.5; DB 2; Length 302;
Best Local Similarity 29.8%; Pred. No. 22;
Matches 14; Conservative 10; Mismatches 12; Indels 11; Gaps 1;

QY   1 MLSNLIRLLNK-----ALRKAHSMVRNFYRGKPVQSOLK 36
    :|:|||||::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~
DB   198 LLDDLRLPNRSDFDKLTGTTSDGLRRFGTWAQSLQYGTPTVGQALR 244

RESULT 14
Q74H89          PRELIMINARY;       PRT;     795 AA.
ID Q74H89
AC Q74H89;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE DNA gyrase, B subunit (EC 5.99.1.3).
GN Name=gyrB; ORFNames=GSU0003;
OS Geobacter sulfurreducens.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfuromonadales;
OC Geobacteraceae; Geobacter.
NCBI_TaxID=33554;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PCA / ATCC 51573;
RX PubMed=146711304; DOI=10.1126/science.1088727;
RA Merhe B.A., NELSON K.E., EISEN J.A., PAULSEN I.T., NELSON W.C.,
RA Heideberg J.F., WU D., WU M., WARD N.L., BEAMAN M.J., DODSON R.J.,
RA Madupu R., BRINKAC L.M., DAUGHERTY S.C., DEBOY R.T., DURKIN A.S.,
RA GWINN M.L., KOLONAY J.F., SULLIVAN S.A., HAFT D.H., SELENGUT J.,
RA DAVIDENCO T.M., ZAFAR N., WHITE O., TRAN B., ROMERO C., FORBERGER H.A.,
```

KW Isomerase. 795 AA; 88786 MW; 339280BFF763576DB CRC64;  
 SQ SEQUENCE 26.7%; Score 55.5; DB 2; Length 795;  
 Query Match 37.5%; Pred. No. 64;  
 Best Local Similarity 15; Conservative 8; Mismatches 10; Indels 7; Gaps 2;  
 Matches 15; Conservative 8; Mismatches 10; Indels 7; Gaps 2;

QY 2 LSNLRILLNKAALKAHT-SWNRFRYGRKPVQSQLPRDL 40  
 DB 656 LGNLRALDQ-----HTLDLTGSYEGLLVESYRKVRDI 689

RESULT 16  
 DPOA\_YEAST STANDARD; PRT; 1468 AA.  
 AC P13382;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE DNA polymerase alpha catalytic subunit (EC 2.7.7.7) (DNA polymerase  
 1).  
 DE Names=POL1; Synonyms=CDL17; OrderedLocusNames=YNL102W; ORFNames=N2181;  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND VARIANT ARG-493.  
 RX MEDLINE=88234507; PubMed=3287376;  
 RA Pizzagalli A., Valasek P., Plevani P., Lucchini G.,  
 RT "DNA polymerase I gene of Saccharomyces cerevisiae: nucleotide  
 sequence, mapping of a temperature-sensitive mutation, and protein  
 RT homology with other DNA polymerases."  
 RL Proc. Natl. Acad. Sci. U.S.A. 85:3772-3776(1988).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C / FY1679;  
 RA Saiz J.E., Butrago M.J., Soler A., del Rey F., Revuelta J.L.;  
 RT "The sequence of a 21.3 kb DNA fragment from the left arm of yeast  
 RT chromosome XIV reveals LEU4, MET4, RAS2, and six new open  
 RT reading frames."  
 RL Yeast 12:403-409(1996).  
 RN [3]  
 RP SEQUENCE OF 1-26 FROM N.A.  
 RC STRAIN=S288C;  
 RX MEDLINE=93188702; PubMed=8446029;  
 RA Mountain H.A., Bystrom A.S., Korch C.;  
 RT "The general amino acid control regulates MET4, which encodes a  
 RT methionine-pathway-specific transcriptional activator of Saccharomycetes  
 RT cerevisiae."  
 RL Mol. Microbiol. 7:215-228(1993).  
 RN [4]  
 RP FUNCTION, AND SUBUNIT.  
 RX MEDLINE=20357110; PubMed=10898792;  
 RA Qi H., Zakian V.A.;  
 RT "The Saccharomyces telomere-binding protein Cdc13p interacts with both  
 RT the catalytic subunit of DNA polymerase alpha and the telomerase-  
 RT associated est1 protein."  
 RL Genes Dev. 14:1777-1788(2000).  
 CC -!- FUNCTION: Polymerase alpha in a complex with DNA primase is a  
 CC replicative polymerase. Has a role in promoting telomere  
 CC replication during interaction with CDC13.  
 CC -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate  
 CC + [DNA] (N).  
 CC -!- SUBUNIT: Interacts with CDC13.  
 CC -!- SUBCELLULAR LOCATION: Nuclear.  
 CC -!- MISCELLANEOUS: In eukaryotes there are five DNA polymerases:  
 CC alpha, beta, gamma, delta, and epsilon which are responsible for  
 CC different reactions of DNA synthesis.  
 CC -!- SIMILARITY: Belongs to the DNA polymerase type-B family.

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL: J03258; AAA34888.1; -;  
 DR EMBL: Z50161; CAA90524.1; -;  
 DR EMBL: Z71378; CAA95978.1; -;  
 DR EMBL: Z12126; CAA78111.1; -;  
 DR PIR: S58250; S58250.  
 DR GerMOnline: 143108.  
 DR SGD: S0005046; POL1.  
 DR GO: GO:0000731; P:DNA repair synthesis; IMP.  
 DR InterPro: IPR006172; DNA\_pol\_B.  
 DR InterPro: IPR006133; DNA\_pol\_B\_exo.  
 DR InterPro: IPR006134; DNA\_pol\_B\_region.  
 DR InterPro: IPR004578; Pol2.  
 DR Pfam: PF00136; DNA\_pol\_B; 1.  
 DR Pfam: PF03104; DNA\_pol\_B\_exo; 1.  
 DR PRINTS: PR00106; DNAPOLB.  
 DR SMART: SM00486; POLB; 1.  
 DR TIGRFAMs: TIGR00592; pol2; 1.  
 DR PROSITE: PS00116; DNA\_POLYMERASE\_B; 1.  
 DR DNA replication; DNA-binding; DNA-directed DNA polymerase;  
 KW Nuclear protein; Transferase.  
 FT DNA BIND 1246 1381 Potential.  
 FT VARIANT 493 493 G -> R (IN TEMPERATURE SENSITIVE MUTANT).  
 FT CONFLICT 759 760 MI -> IV (in Ref. 1).  
 SQ SEQUENCE 1468 AA; 16808 MW; 50C9032DBE9585AE CRC64;  
 Query Match 26.7%; Score 55.5; DB 1; Length 1468;  
 Best Local Similarity 29.4%; Pred. No. 1.3e+02;  
 Matches 15; Conservative 4; Mismatches 21; Indels 11; Gaps 1;  
 QY 2 LSNLRILL-----NKAALKAHTSMVNRFRYGRKPVQSQLPRDL 41  
 DB 533 MPNLRCLSLSIQTLMPNKPENKQIVSITLSAYNRITSLDSPENIKRPDDL 583

RESULT 17  
 Q6P0K2 PRELIMINARY; PRT; 258 AA.  
 ID Q6P0K2  
 AC Q6P0K2;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Hypothetical protein zgc:77155.  
 GN Name=zgc:77155;  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7957;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Faxner A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Greenwood J., Schmutz J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RA Strausberg R.;  
 RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC065586; AAH65586.1; -;  
 DR InterPro: IPR000504; RNA\_rec\_not.  
 DR InterPro: IPR001878; Znf\_CCHC.  
 DR Pfam: PF00076; RRM\_1; 1.  
 DR Pfam: PF00098; zf-CCHC; 1.  
 DR PRINTS: PR00939; C2HCZNFINGER.  
 DR SMART: SM00360; RRM; 1.  
 DR SMART: SM00343; Znf\_C2HC; 1.  
 DR PROSITE: PS0102; RRM; 1.  
 DR PROSITE: PS0158; ZF\_CCHC; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 258 AA; 29630 MW; 9ABDF16B9DA007B4 CRC64;  
 Query Match 26.4%; Score 55; DB 2; Length 258;  
 Best Local Similarity 47.6%; Pred. No. 21;  
 Matches 10; Conservative 2; Mismatches 9; Indels 0; Gaps 0;  
 QY 21 MVRNFRYKPKVQSQLKPRDLC 41  
 DB 88 MSKSRYGPRSRQDFDNDRC 108  
 RESULT 18  
 AAH65586 PRELIMINARY; PRT; 258 AA.  
 AC AAH65586;  
 DT 24-MAY-2004 (TrEMBLrel. 27, Created)  
 DT 24-MAY-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 24-MAY-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Hypothetical protein zgc:77155.  
 GN ZGC:77155.  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahney J., Heiton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.

RC TISSUE=Kidney;  
 RA Strausberg R.;  
 RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC065586; AAH65586.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 258 AA; 29630 MW; 9ABDF16B9DA007B4 CRC64;  
 Query Match 26.4%; Score 55; DB 2; Length 258;  
 Best Local Similarity 47.6%; Pred. No. 21;  
 Matches 10; Conservative 2; Mismatches 9; Indels 0; Gaps 0;  
 QY 21 MVRNFRYKPKVQSQLKPRDLC 41  
 DB 88 MSKSRYGPRSRQDFDNDRC 108  
 RESULT 19  
 Q6CJV5 PRELIMINARY; PRT; 338 AA.  
 AC Q6CJV5;  
 DT 01-OCT-2004 (TrEMBLrel. 28, Created)  
 DT 01-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
 DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE Strain NRRL Y-1140 chromosome P of strain NRRL Y-1140 of Kluyveromyces  
 DE lactis.  
 GN ORFNames=KLIA0F15664g;  
 OS Kluyveromyces lactis (Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.  
 OX NCBI\_TaxID=28985;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NRRL Y-1140;  
 RG GENOLEVURES;  
 RA Dujon B., Sherman D., Fischer G., Durrans P., Casaregola S.,  
 RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla E.,  
 RA Goffard N., Frangeul L., Algic M., Anchoard V., Babour A., Barbe V.,  
 RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,  
 RA Boisrame A., Boyer J., Cattolico L., Confanioli F., de Daruvar A.,  
 RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,  
 RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,  
 RA Kerrast A., Keszul R., Lemaire M., Lesur I., Ma L., Muller H.,  
 RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,  
 RA Pellenn S., Potier S., Richard G.F., Straub M.L., Suleau A.,  
 RA Swennene D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,  
 RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,  
 RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,  
 RA Wincker P., Souciet J.L.;  
 RT "Genome evolution in yeasts.";  
 RL Nature 430:35-44 (2004).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NRRL Y-1140;  
 RA Genoscope;  
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: CR382126; CAG98492.1; -;  
 SQ SEQUENCE 338 AA; 38355 MW; 4BF496B7448583B2 CRC64;  
 Query Match 26.4%; Score 55; DB 2; Length 338;  
 Best Local Similarity 40.0%; Pred. No. 29;  
 Matches 12; Conservative 6; Mismatches 12; Indels 0; Gaps 0;  
 QY 10 NKAALKAHTSMVRNFRYKPKVQSQLKPRD 39  
 DB 106 NIAALKNLKTSAILSFSVAGSLQKEIKPRD 135  
 RESULT 20  
 Q9JUG9 PRELIMINARY; PRT; 138 AA.  
 ID Q9JUG9  
 AC Q9JUG9;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Hypothetical protein NMA1316.  
GN OrderedLocusNames=NMA1316;  
OS Neisseria meningitidis (serogroup A).  
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;  
OC Neisseriaceae; Neisseria.  
OX NCBI\_TaxID=65699;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=22491 / Serogroup A / Serotype 4A;  
RX MEDLINE=20222556; PubMed=10761919; DOI=10.1038/35006655;  
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.M.,  
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,  
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holtroyd S.,  
RA Jagels K., Leather S., Moule S., Mungall K.L., Quail M.A.,  
RA Rajandream M.A., Rutherford S., Mungall K.L., Simmonds M., Skelton J.,  
RA Whitehead S., Spratt B.G., Barrall B.G.;  
RT "Complete DNA sequence of a serogroup A strain of Neisseria  
meningitidis 22491".  
RL Nature 404:502-506(2000).  
DR EMBL; AL162755; CAB84568.1; -;  
DR PIR; F81900; F81900.  
DR InterPro; IPR007110; Ig-like.  
DR IntraPro; IPR006522; Tail comp.S.  
DR Pfam; PF05069; Phage\_tail\_S; 1.  
DR TIGRFAMs; TIGR01635; tail\_comp.S; 1.  
KW Complete proteome; Hypothetical protein.  
SQ SEQUENCE 138 AA; 15738 MW; FC0BB880236F2113 CRC64;  
  
Query Match 26.2%; Score 54.5; DB 2; Length 138;  
Best Local Similarity 42.9%; Pred. No. 13; Mismatches 15; Indels 1; Gaps 1;  
Matches 15; Conservative 4; Mismatches 15; Indels 1; Gaps 1;  
  
QY 6 RILLNKAALRKAKHTSMVRNFRY-GKPVQSQLKPRD 39  
DB 26 RYLLMRRLSETMTAVKLNFRYAGRPKWLGLKYRD 60  
  
RESULT 21  
ID R2 LOTJA STANDARD; PRT; 236 AA.  
AC Q9BES6;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Chloroplast 30S ribosomal protein S2.  
GN Name-rps2;  
OS Lotus japonicus.  
OG Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae; Lotus.  
OX NCBI\_TaxID=34305;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Accession MG-20;  
RX MEDLINE=21082929; PubMed=11214967;  
RA Kato T., Kaneko T., Sato S., Nakamura Y., Tabata S.;  
RT "Complete structure of the chloroplast genome of a legume, Lotus  
japonicus".  
RL DNA Res. 7:323-330(2000).  
CC -1- SIMILARITY: Belongs to the S2P family of ribosomal proteins.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.ebi-sib.ch/announce/>  
CC or send an email to [license@ebi-sib.ch](mailto:license@ebi-sib.ch)).  
CC EMBL; AP002983; BAB33197.1; -;  
DR HAMAP; MF\_00291; -; 1.

DR InterPro; IPR001865; Ribosomal\_S2.  
DR InterPro; IPR005706; Ribosomal\_S2\_b/o.  
DR Pfam; PF00318; Ribosomal\_S2; 1.  
DR PRINTS; PRO0395; RIBOSOMALS2.  
DR TIGRFAMs; TIGR01011; rpsB\_bact; 1.  
DR PROSITE; PS00962; RIBOSOMAL\_S2\_1; 1.  
DR PROSITE; PS00963; RIBOSOMAL\_S2\_2; 1.  
KW Chloroplast; Ribosomal protein.  
SQ SEQUENCE 236 AA; 26982 MW; CE9E238572325586 CRC64;  
  
Query Match 26.0%; Score 54; DB 1; Length 236;  
Best Local Similarity 38.5%; Pred. No. 27;  
Matches 10; Conservative 10; Mismatches 4; Indels 2; Gaps 1;  
  
QY 2 LNLRLILNKA--ALRKAHTSMVRNF 25  
DB 211 IASIRLILNKLVAICEGHSYIRNF 236  
  
RESULT 22  
ID Q73ZD6 PRELIMINARY; PRT; 271 AA.  
AC Q73ZD6;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Hypothetical protein.  
GN OrderedLocusNames=MAP1667;  
OS Mycobacterium paratuberculosis.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;  
OC Mycobacterium avium complex (MAC).  
OX NCBI\_TaxID=1770;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=k10;  
RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;  
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AE017233; AAS03984.1; -;  
DR InterPro; IPR006311; Tat.  
DR TIGRFAMs; TIGR01409; Tat signal seq; 1.  
KW Complete proteome; Hypothetical protein.  
SQ SEQUENCE 271 AA; 28478 MW; 3DAFFC3DD8019D71 CRC64;  
  
Query Match 26.0%; Score 54; DB 2; Length 271;  
Best Local Similarity 61.1%; Pred. No. 32;  
Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
  
QY 13 ALRKAHTSMVRNFRYCKP 30  
DB 101 ALRAAGLHIVSNFQYCKP 118  
  
RESULT 23  
ID AAS03984 PRELIMINARY; PRT; 271 AA.  
AC AAS03984;  
DT 02-MAR-2004 (TrEMBLrel. 27, Created)  
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)  
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)  
DE Hypothetical protein.  
GN MAP1667.  
OS Mycobacterium paratuberculosis.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1770;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=k10;  
RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;  
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AE017233; AAS03984.1; -;  
KW Hypothetical protein.

```
SQ SEQUENCE 271 AA; 28478 MW; 3DAFFC3DD8019D71 CRC64;
Query Match .26.0%; Score 54; DB 2; Length 271;
Best Local Similarity 61.1%; Pred. No. 32;
Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 13 ALRKAHTSMVRNFRYKGP 30
||| :||| :|||
Db 101 ALRAAGLHIVSNFYQYKGP 118

RESULT 24
Q8AVS9 PRELIMINARY; PRT; 457 AA.
ID Q8AVS9
AC Q8AVS9;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE MGC53557 protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OC NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Rappaport M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
Rosa S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahy J., Helton E., Kettman M., Madan A.C., Rodriguez S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
Krzyszewski M.I., Skalska U., Smalley D.E., Schnerch A., Schein J.E.,
Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
initiative";
RL Dev. Dyn. 225:384-391(2002).
[3]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 3 WD repeats.
DR EMBL; BC041284; AAH41284.1; -
DR InterPro; IPR001680; WD40.
DR InterPro; IPR011046; WD40_like.
DR Pfam; PF00400; WD40; 3.
DR ProDom; PD000018; WD40; 1.
DR SMART; SM00320; WD40; 5.
DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_1.
DR PROSITE; PS00082; WD_REPEATS_2; 2.
DR PROSITE; PS00294; WD_REPEATS_REGION; 1.
```

```
KW Repeat; WD repeat.
SQ SEQUENCE 457 AA; 51021 MW; EE44FE225DD0039B CRC64;
Query Match .26.0%; Score 54; DB 2; Length 457;
Best Local Similarity 46.4%; Pred. No. 57;
Matches 13; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

QY 2 LSNRLILLNKAALRKAHTSMVRNFRYK 29
||| :||| :||| :|||
Db 135 LWDLRKLNKCACTLHGTHTSWKNIEYDK 162

RESULT 25
A37C_DROLE STANDARD; PRT; 544 AA.
ID A37C_DROLE
AC Q96570;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Anon-37Cs protein.
GN Names=anon-37Cs; Synonyms=Cs; ORFNames=CG10561;
OS Drosophila lebanonensis (fruit fly) (Scaptodrosophila lebanonensis).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Scaptodrosophila.
OC NCBI_TaxID=7225;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Beitut;
RX MEDLINE=99250256; PubMed=10231575;
RA Tatarenkov A., Saez A.G., Ayala F.J.;
RT "A compact gene cluster in Drosophila: the unrelated Cs gene is
compressed between duplicated amd and Ddc.";
RL Gene 231:111-120(1999).
CC -1- FUNCTION: Has a nonvital function (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC EMBL; AF091329; AAC67584.1; -
DR FlyBase; FBgn0025668; Dlebl\CG10561.
DR GO; GO:0005737; Cytoplasm; ISS.
DR InterPro; IPR002937; Amino_oxidase.
DR InterPro; IPR000205; NAD_BS.
DR Pfam; PF01593; Amino_oxidase; 1.
SQ SEQUENCE 544 AA; 61026 MW; CF003E2CBB7D6DBE CRC64;
Query Match .26.0%; Score 54; DB 1; Length 544;
Best Local Similarity 37.1%; Pred. No. 69;
Matches 13; Conservative 4; Mismatches 18; Indels 0; Gaps 0;

QY 1 MLNRLILLNKAALRKAHTSMVRNFRYKPVQSOL 35
||| :||| :||| :||| :|||
Db 332 VLKNFSAILFKPALPLEKQAIRNLGYGNPVKIYL 366

RESULT 26
Q52393 PRELIMINARY; PRT; 769 AA.
ID Q52393
AC Q52393;
DT 01-NOV-1996 (T-EMBLrel. 01, Created)
DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (T-EMBLrel. 22, Last annotation update)
DE HrpY protein.
GN Names=hrpY;
OS Pseudomonas syringae (pv. phaseolicola).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
```





DR GO:0006118; P-electron transport; IEA.  
DR InterPro: IPR006090; Acyl-CoA dh.C  
DR InterPro: IPR006091; Acyl-CoA dh.M  
DR InterPro: IPR009075; AcylCoADH\_C-like.  
DR InterPro: IPR009100; AcylCoA dehyd\_NM.  
DR Pfam: PF00441; Acyl-CoA dh; i.  
DR Pfam: PF02770; Acyl-CoA dh.M; l.  
KW FAD; Flavoprotein; Oxidoreductase.  
SQ SEQUENCE 383 AA; 42742 MW; E20129A4F567A8C2 CRC64;

Query Match            25.7%; Score 53.5; DB 2; Length 383;  
Best Local Similarity     42.9%; Pred.No.55;  
Matches      15; Conservative       3; Mismatches    14; Indels          3; Gaps           1;

OY            2 LSNLRLLNKAALRKAAHTSMVNFYKGVQSQLK 36  
DB            :|::||| || | | | | | | | | | |  
ID Q6LGI0 PRELIMINARY; PRT; 236 AA.

AC Q7SGU4 PRELIMINARY; PRT; 1113 AA.  
DT 01-MAR-2004 (TREMBLrel. 26, Created)  
DT 01-JUN-1998 (TREMBLrel. 26, Last sequence update)  
DE Hypothetical protein BBG23.  
GN Name=NCU03225.1;  
OS Neurospora crassa.  
OC Eukaryota; Fungi; Ascomycota; Peizomycotina; Sordariomycetes;  
OX NCBI\_TaxID=5141;  
RN [1] \_TaxID=5141;  
RP SEQUENCE FROM N.A.  
RC STRAIN=OR74A;  
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,  
RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,  
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,  
RA Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,  
RA Sellitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,  
RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,  
RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnere S.,  
RA Kamal M., Kamysysselis M., Maucelli E., Bielke C., Rudd S., Frishman D.,  
RA Krystofova S., Rasmussen C., Metznerberg R.L., Perkins D.D., Kroken S.,  
RA Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Osmani S.A.,  
RA DeSouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,  
RA Varden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,  
RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,  
RA Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birn B.;  
RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";  
RL Nature 0-0(2003).  
CC -! CAUTION: The sequence shown here is derived from an  
CC EMBL/GenBank/DDSB whole genome shotgun (WGS) entry which is preliminary data.  
DR EMBL; AABXO100007; EAA36089.1; -.  
RW Hypothetical protein.  
SQ SEQUENCE 1113 AA; 125888 MW; 7623048B96335024 CRC64;

Query Match            25.7%; Score 53.5; DB 2; Length 1113;  
Best Local Similarity     36.2%; Pred.No.1.8e+02;  
Matches      17; Conservative       8; Mismatches    13; Indels          9; Gaps           2;

OY            2 LSNR-----ILLNKALKRKAHTSMVFN----FRYGKPVQSOLKPD 39  
DB            |:|:||| | | | | | | | | | | | | | |:  
ID Q6LGIO PRELIMINARY; PRT; 236 AA.

ID Q6LGIO PRELIMINARY; PRT; 236 AA.  
AC Q6LGIO;  
DT 05-JUL-2004 (TREMBLrel. 27, Created)  
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)







```
DR InterPro; IPR000178; IPF2.
```

Query Match  
Best Local Similarity    25.5%; Score 53; DB 1; Length 597;  
Matches    13; Conservative    12; Mismatches    11; Indels         6; Gaps      2;

DY

QY                  3 SNLRILNKAALRKAHTSMWRNFYRGKPVQS---LKPRLD 40

DB                 255 ATDLVLYDGTLKKGDVTWIGSL--GEPIQTKVALLKPREL 294

RESULT 40

Q99A14 PRELIMINARY; PRT; 140 AA.

ID Q99A14

AC Q99A14

DT 01-JUN-2001 (TrEMBLrel. 17, Created)

DI DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DE E2 glycoprotein (Fragment).

OS Bovine viral diarrhea virus strain 11203/98.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

CC Pestivirus.

NCBI\_TaxID=145228;

[1]

RN SEQUENCE FROM N.A.

RP STRAIN=11203/98;

RA Tajima M.;

RL Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.

[2]

RN SEQUENCE FROM N.A.

RP STRAIN=11203/98;

RC Greiser-Wilke I.M.;

RL Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.

EMBL; AJ303004; CAC24802.1; -.

FT NON TER 1

FT NON TER 140

SQ SEQUENCE 140 AA; 15335 MW; CS93027AAB37E0FC CRC64;

Query Match                  25.2%; Score 52.5; DB 2; Length 140;  
Best Local Similarity        35.3%; Pred.No.25;  
Matches    11; Conservative    7; Mismatches    7; Indels         3; Gaps      1;

QY                  7 ILNKAALRKAHTSMWRNFYRGKPVQSQ 34

DB                 2 VLNVKATL---HTAVVITYKEASFPFSR 26

Search completed: December 18, 2004, 02:49:59  
Job time : 126.488 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 18, 2004, 02:21:06 ; Search time 115.179 Seconds  
(without alignments)  
133.926 Million cell updates/sec

Title: US-08-765-244-22

Perfect score: 217

Sequence: 1 MLSNRIILLKAAALRAHTS.....NFRYKPVQSVQLKPRDLIC 43

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- A\_Geneseq\_23Sep04:\*
- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	217	100.0	43	2 AAR90584	Aar90584 Rat ornit
2	199	91.7	354	7 ADD47079	Add47079 Rat Prote
3	199	91.7	354	7 ADD48634	Add48634 Rat Prote
4	159	73.3	32	5 ABG30857	Abg30857 Rat ornit
5	141	65.0	354	7 ADD48636	Add48636 Human Pro
6	141	65.0	354	7 ADD47081	Add47081 Human Pro
7	109	50.2	358	3 AAB18445	Aab18445 Protein e
8	108	49.8	32	5 ABG30856	Abg30856 Human orn
9	105	48.4	32	2 AAR48224	Aar48224 ORC pepti
10	92	42.4	32	2 AAR48226	Aar48226 Mitochond
11	92	42.4	32	2 AAW96358	Aaw96358 Mitochond
12	92	42.4	32	5 ABG92993	Abg92993 Localisat
13	92	42.4	32	6 ABP56588	Abp56588 Mitochond
14	79.5	36.6	31	3 AAB22835	Aab22835 Mitochond
15	79.5	36.6	31	3 AAB15704	Aab15704 Mitochond
16	56.5	26.0	149	4 AAU23468	Aau23468 Novel hum
17	54	24.9	866	4 ABB57741	Abb57741 Drosophil
18	53.5	24.7	755	4 ABBS9259	Abbs9259 Drosophil
19	53	24.4	457	7 ABO84066	Abo84066 Pseudomon
20	53	24.4	593	6 ABUS0395	Abus0395 Protein e
21	52	24.0	165	2 AAY04933	Aay04933 Mycobacte
22	52	24.0	218	7 ABO69307	Abo69307 Pseudomon
23	52	24.0	774	4 ABG29304	Abg29304 Novel hum
24	52	24.0	1124	4 ABB59241	Abb59241 Drosophil
25	51.5	23.7	138	3 AAY75465	Aay75465 Neisseria

26	51.5	23.7	291	6 ABM68583	Abm68583 Photorhab
27	51	23.5	65	7 ADH87578	Adh87578 Enterococ
28	51	23.5	598	6 ABU45406	Abu45406 Protein e
29	51	23.5	705	6 ABR53783	AbR53783 Protein s
30	51	23.5	705	7 ADK63830	Adk63830 Disease t
31	51	23.5	1169	2 AAR96126	Aar96126 Bacillus
32	50.5	23.3	138	3 AAY75466	Aay75466 Neisseria
33	50.5	23.3	296	4 ABB64170	Abb64170 Drosophil
34	50.5	23.3	533	6 ABM67661	Abm67661 Photorhab
35	50.5	23.3	2379	5 ABP62760	Abp62760 S. roseos
36	50.5	23.3	2379	7 ADJ72172	Adj72172 Streptomy
37	50	23.0	564	5 ABP73265	Abp73265 Candida a
38	50	23.0	746	4 AAB93938	Aab93938 Human pro
39	50	23.0	975	4 AAB94042	Aab94042 Human pro
40	50	23.0	975	4 AAO18173	Aao18173 Human hyd
41	50	23.0	975	8 ADK60418	Adk60418 Angiogene
42	50	23.0	975	8 ADK60194	Adk60194 Angiogene
43	50	23.0	975	8 ADK60495	Adk60495 Angiogene
44	50	23.0	975	8 ADK60719	Adk60719 Angiogene
45	50	23.0	975	8 ADP73118	Adp73118 Angiogene

ALIGNMENTS

RESULT 1

AAR90584

ID AAR90584 standard; protein; 43 AA.

XX AAR90584;

DT 25-MAR-2003 (revised)

DT 31-OCT-1996 (first entry)

DE DE Rat ornithine transcarbamylase signal peptide.

KW promoter; peptide-nucleic acid; cyclised; gene therapy; target;

KW site-directed mutagenesis; introduction; protein transport.

XX Synthetic.

PN DE19520815-A1.

PD 21-DEC-1995.

PF 11-JUN-1995; 95DE-01020815.

PR 16-JUN-1994; 94DE-04421079.

XX (SEIB/) SEIBEL P.

PI Seibel P, Seibel A;

DR WPI; 1996-041226/05.

PT Replicable and transcriptionally active plasmid carrying signal peptide  
for specific target - useful for site directed mutagenesis and molecular  
therapy of genetic diseases.

XX Disclosure; Col 11; 24pp; German.

XX Two modified oligonucleotides (introducing PstI and XhoI sites) were used  
to amplify a region of the human mitochondrial (mt) genome contg. the  
light strand promoter, mtDNA ori of the heavy strand, CSB's ("conserved  
sequence blocks") and a regulation site for DNA replication. Behind this  
fragment (5' direction) a synthetic multiple cloning site was introduced,  
generating a product with overhaign ends. The synthetic region also  
introduced a bidirectional mt transcription termination sequence. The  
amplification product, synthetic fragment and pluescript were ligated  
and recombinant plasmid 1 (AAT12315) was produced. Human mt 16S rRNA  
(differing from the native RNA only in having a modified nucleotide) was  
isolated by PCR from chloramphenicol resistant HeLa cells and inserted  
into plasmid 1 to form plasmid 2 (AAT12316). The cloned insert was

CC isolated as a Bsal fragment and cyclised using hairpin loop  
 CC oligonucleotides, one of which carried the required signal peptide (the  
 CC present sequence). The cyclised product was purified by treatment with  
 CC exonuclease III. In a modification, the signal peptide was attached after  
 CC cyclisation. The new plasmids were able to impart chloramphenicol  
 CC resistance to otherwise sensitive B lymphocytes and fibroblasts. Similar  
 CC plasmids without a signal peptide could not do this. The plasmids can be  
 CC introduced into eukaryotic cells, esp. for site-directed mutagenesis or  
 CC molecular therapy of genetic diseases, targeting nucleic acid in cells  
 CC or their organelles via the protein transport route. (Updated on 25-MAR-  
 CC 2003 to correct PR field.)

XX Sequence 43 AA;

Query Match 100.0%; Score 217; DB 2; Length 43;  
 Best Local Similarity 100.0%; Pred. No. 6.2e-26;  
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLSNRLILNKALRKAHTSMVRNFRYKPVQSQVQLKPRDLC 43  
 Db 1 MLSNRLILNKALRKAHTSMVRNFRYKPVQSQVQLKPRDLC 43

RESULT 2  
 ADD47079  
 ID ADD47079 standard; protein; 354 AA.

XX ADD47079;

XX 29-JAN-2004 (first entry)

DE Rat Protein AAA41767, SEQ ID NO 12767.

Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
 Chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

OS Rattus norvegicus.

PN WO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

XX (GEHO ) GEN HOSPITAL CORP.

PA (FARB ) BAYER AG.

PI Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

DR GENBANK; AAA41767.

XX New composition comprising two or more isolated polypeptides, useful for

PT preparing a medicament for treating pain in an animal.

XX Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a

CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 354 AA;

Query Match 91.7%; Score 199; DB 7; Length 354;  
 Best Local Similarity 97.6%; Pred. No. 5.3e-22;  
 Matches 41; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MLSNRLILNKALRKAHTSMVRNFRYKPVQSQVQLKPRDL 42  
 Db 1 MLSNRLILNKALRKAHTSMVRNFRYKPVQSQVQLKGRDL 42

RESULT 3

ADD48634

ID ADD48634 standard; protein; 354 AA.

XX ADD48634;

DT 29-JAN-2004 (first entry)

XX Rat Protein OMRT, SEQ ID NO 14340.

Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
 Chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

OS Rattus norvegicus.

PN WO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

XX (GEHO ) GEN HOSPITAL CORP.

PA (FARB ) BAYER AG.

PI Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

XX New composition comprising two or more isolated polypeptides, useful for

PT preparing a medicament for treating pain in an animal.

XX Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC that is differentially expressed in neuronal tissue of a first animal



CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a rat protein (shown in table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 354 AA;  
SQ

Query Match 91.7%; Score 199; DB 7; Length 354;  
Best Local Similarity 97.6%; Pred. No. 5.3e-22;  
Matches 41; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLSNRLILNKAAALRAKHAHTSMVRNFRYKGPVQSVQVQLKPRDL 42  
DB 1 MLSNRLILNKAAALRAKHAHTSMVRNFRYKGPVQSVQVQLKGRDL 42  
|||||

RESULT 4  
ABG30857  
ID ABG30857 standard; peptide; 32 AA.  
XX  
AC ABG30857;  
XX  
DT 07-OCT-2002 (first entry)  
XX  
DE Rat ornithine transcarbamylase signal peptide.  
XX  
KW Rat; signal peptide; ornithine transcarbamylase; MOT; recombinant vector;  
KW fusion protein; extranuclear gene.  
XX  
OS Rattus sp.  
XX  
PN JP2002176988-A.  
PD  
PP 25-JUN-2002.  
XX  
PF 14-DEC-2000; 2000JP-00380975.  
XX  
PR 14-DEC-2000; 2000JP-00380975.  
XX  
PA (TANA/) TANAKA M.  
PA (GIFU-) ZH GIFUKEN KOKUSAI BIO KENTUJO.  
PA (OYOS-) ZH OYO SEIKAGAKU KENKUYUSHO.  
XX  
DR WPI; 2002-569946/61.  
DR N-PSDB; ABX8419.  
XX  
PT A recombinant vector for expressing a fused protein, useful for  
PT decomposing an extranuclear gene of a nonhuman organism.  
XX  
PS Disclosure; Page 3; 15pp; Japanese.  
XX

CC The invention relates to a recombinant vector for expressing a fused  
CC protein containing a fused gene in which a base sequence defining a  
CC transfer signal peptide to small cellular organs having an extranuclear  
CC gene is combined with a base sequence defining the amino acid sequence of  
CC a restriction enzyme recognising a defined base sequence. The vector is  
CC used for decomposing an extranuclear gene of a nonhuman organism. The  
CC present sequence represents the signal peptide of rat mitochondrial  
CC ornithine transcarbamylase (MOT) which may be used in the vector of the

CC invention  
XX Sequence 32 AA;  
SQ

Query Match 73.3%; Score 159; DB 5; Length 32;  
Best Local Similarity 100.0%; Pred. No. 4.7e-17;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLSNRLILNKAAALRAKHAHTSMVRNFRYKGPVQ 32  
DB 1 MLSNRLILNKAAALRAKHAHTSMVRNFRYKGPVQ 32  
|||||

RESULT 5  
ADD48636  
ID ADD48636 standard; protein; 354 AA.  
XX  
AC ADD48636;  
XX  
DT 29-JAN-2004 (first entry)  
XX  
DE Human Protein P00480, SEQ ID NO 14342.  
XX  
KW Human; pain; neuronal tissue; gene therapy;  
KW spinal segmental nerve injury; Chronic constriction injury; CCI;  
KW spared nerve injury; SNI; Chung.  
XX  
OS Homo sapiens.  
XX  
PN WO2003016475-A2.  
XX  
PD 27-FEB-2003.  
XX  
PF 14-AUG-2002; 2002WO-US025765.  
XX  
PR 14-AUG-2001; 2001US-0312147P.  
PR 01-NOV-2001; 2001US-0346382P.  
PR 26-NOV-2001; 2001US-0333347P.  
XX  
PA (GEHO) GEN HOSPITAL CORP.  
PA (FARB) BAYER AG.  
PI Woolf C, D'urso D, Befort K, Costigan M;  
XX  
DR WPI; 2003-268312/26.  
DR GENBANK; P00480.  
XX  
PT New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.  
XX  
PS Claim 1; Page; 1017pp; English.  
XX

CC The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction











CC N-end rule pathway. This method for regulating protein stability allows  
CC removal of the antibody after it has bound to its target antigen.  
CC Stabilised recombinant proteins may be used in gene therapy for the  
CC treatment of disorders such as Alzheimer's disease

XX SQ Sequence 31 AA;

Query Match 36.6%; Score 79.5; DB 3; Length 31;

Best Local Similarity 61.3%; Pred. No. 0.00011;

Matches 19; Conservative 1; Mismatches 10; Indels 1; Gaps 1;

QY 1 MLSNLRILNKAALRKAHTSMVNFYRGKPV 31

DB 1 MLENLR-XLNNAAFRGHFWNFVRCGGL 30

RESULT 16

AAU23468

ID AAU23468 standard; protein; 149 AA.

XX AC AAU23468;

XX DT 18-DEC-2001 (first entry)

DE DE Novel human enzyme polypeptide #554.

XX KW Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;

KW ligase; hyperproliferative disorder; immunodeficiency disorder;

KW autoimmune disorder; neurological disorder; metabolic disorder;

KW inflammatory disorder; cardiovascular disorder; reproductive disorder;

KW blood-related disorder; infectious disorder; cytostatic; anti arthritic;

KW nephrotropic; anticoagulant.

XX OS Homo sapiens.

XX PN WO200155301-A2.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US0001239.

XX PR 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 11-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 18-AUG-2000; 2000US-0225759P.

PR 18-AUG-2000; 2000US-0226279P.

PR 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226868P.

PR 22-AUG-2000; 2000US-0227182P.

PR 23-AUG-2000; 2000US-0227009P.

PR 30-AUG-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.

PR 01-SEP-2000; 2000US-0229343P.

PR 01-SEP-2000; 2000US-0229344P.

PR 01-SEP-2000; 2000US-0229345P.

PR 05-SEP-2000; 2000US-0229509P.

PR 05-SEP-2000; 2000US-0229513P.

PR 06-SEP-2000; 2000US-0230437P.

PR 06-SEP-2000; 2000US-0230438P.

PR 08-SEP-2000; 2000US-0231242P.

PR 08-SEP-2000; 2000US-0231243P.

PR 08-SEP-2000; 2000US-0231244P.

PR 08-SEP-2000; 2000US-0231413P.

PR 08-SEP-2000; 2000US-0231414P.

PR 08-SEP-2000; 2000US-0232080P.

PR 08-SEP-2000; 2000US-0232081P.

PR 12-SEP-2000; 2000US-0231968P.

PR 14-SEP-2000; 2000US-0232397P.

PR 14-SEP-2000; 2000US-0232398P.

PR 14-SEP-2000; 2000US-0232399P.

PR 14-SEP-2000; 2000US-0232400P.

PR 14-SEP-2000; 2000US-0232401P.

PR 14-SEP-2000; 2000US-0233063P.

PR 14-SEP-2000; 2000US-0233064P.

PR 14-SEP-2000; 2000US-0233065P.

PR 21-SEP-2000; 2000US-0234223P.

PR 21-SEP-2000; 2000US-0234274P.

PR 25-SEP-2000; 2000US-0234997P.

PR 25-SEP-2000; 2000US-0234998P.

PR 26-SEP-2000; 2000US-0235484P.

PR 27-SEP-2000; 2000US-0235834P.

PR 27-SEP-2000; 2000US-0235836P.

PR 29-SEP-2000; 2000US-0236327P.

PR 29-SEP-2000; 2000US-0236367P.

PR 29-SEP-2000; 2000US-0236368P.

PR 29-SEP-2000; 2000US-0236369P.

PR 29-SEP-2000; 2000US-0236370P.

PR 02-OCT-2000; 2000US-0236802P.

PR 02-OCT-2000; 2000US-0237037P.

PR 02-OCT-2000; 2000US-0237038P.

PR 02-OCT-2000; 2000US-0237039P.

PR 02-OCT-2000; 2000US-0237040P.

PR 13-OCT-2000; 2000US-0239355P.

PR 13-OCT-2000; 2000US-0239377P.

PR 20-OCT-2000; 2000US-0240960P.

PR 20-OCT-2000; 2000US-0241221P.

PR 20-OCT-2000; 2000US-0241785P.

PR 20-OCT-2000; 2000US-0241786P.

PR 20-OCT-2000; 2000US-0241787P.

PR 20-OCT-2000; 2000US-0241808P.

PR 20-OCT-2000; 2000US-0241809P.

PR 20-OCT-2000; 2000US-0241826P.

PR 01-NOV-2000; 2000US-0244617P.

PR 08-NOV-2000; 2000US-0246474P.

PR 08-NOV-2000; 2000US-0246475P.

PR 08-NOV-2000; 2000US-0246476P.

PR 08-NOV-2000; 2000US-0246477P.

PR 08-NOV-2000; 2000US-0246478P.

PR 08-NOV-2000; 2000US-0246523P.

PR 08-NOV-2000; 2000US-0246524P.

PR 08-NOV-2000; 2000US-0246525P.

PR 08-NOV-2000; 2000US-0246526P.

PR 08-NOV-2000; 2000US-0246527P.

PR 08-NOV-2000; 2000US-0246528P.

PR 08-NOV-2000; 2000US-0246532P.

PR 08-NOV-2000; 2000US-0246609P.

PR 08-NOV-2000; 2000US-0246610P.

PR 08-NOV-2000; 2000US-0246611P.

PR 08-NOV-2000; 2000US-0246613P.

PR 17-NOV-2000; 2000US-0249207P.

```
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-465566/50.
DR N-PSDB; AAS41338.
XX
XX Novel polypeptides and polynucleotides useful for diagnosing, preventing,
PT treating neural, immune system, muscular, reproductive, pulmonary,
PT cardiovascular, renal, proliferative disorders and cancerous diseases.
XX
XX Claim 11; SEQ ID NO 1464; 1180pp; English.
XX
XX The present invention relates to the isolation of novel human enzyme
CC polypeptides, and the cDNA (AAS40785-AAS41684) and genomic sequences
CC encoding them. The enzyme polypeptides of the invention may comprise the
CC functional classes of oxidoreductases, transferases, hydrolases, lyases,
CC isomerases or ligases. The sequences of the invention are useful in the
CC diagnosis, treatment, prevention and/or prognosis of a wide range of
CC disorders including hyperproliferative disorders (e.g. cancer),
CC immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g.
CC arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic
CC disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma),
CC cardiovascular disorders (e.g. atherosclerosis), blood-related disorders
CC (e.g. haemophilia), reproductive disorders (e.g. infertility) and
CC infectious disorders (e.g. influenza). The polynucleotides of the
CC invention can also be used in gene therapy. AAU22915-AAU23814 represent
CC the novel human enzyme polypeptides of the invention. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 149 AA;
Query Match 26.0%; Score 56.5; DB 4; Length 149;
Best Local Similarity 35.3%; Pred. No. 2.9;
Matches 18; Conservative 7; Mismatches 11; Indels 15; Gaps 3;
QY 5 LRLLNKK-----NALRKAAHTSMVNRFRYG-----KPVQSQVOL-KPR 40
Db 52 LRMTLKESAMIAKSRKPHATMIRNKAHYGLHAGWSLWLPVESALQSHQPR 102
17-NOV-2000; 2000US-0249208P.
17-NOV-2000; 2000US-0249209P.
17-NOV-2000; 2000US-0249210P.
17-NOV-2000; 2000US-0249211P.
17-NOV-2000; 2000US-0249212P.
17-NOV-2000; 2000US-0249213P.
17-NOV-2000; 2000US-0249214P.
17-NOV-2000; 2000US-0249215P.
17-NOV-2000; 2000US-0249216P.
17-NOV-2000; 2000US-0249217P.
17-NOV-2000; 2000US-0249218P.
17-NOV-2000; 2000US-0249244P.
17-NOV-2000; 2000US-0249245P.
17-NOV-2000; 2000US-0249264P.
17-NOV-2000; 2000US-0249265P.
17-NOV-2000; 2000US-0249297P.
17-NOV-2000; 2000US-0249299P.
17-NOV-2000; 2000US-0249300P.
01-DEC-2000; 2000US-0250160P.
01-DEC-2000; 2000US-0250391P.
05-DEC-2000; 2000US-0251030P.
05-DEC-2000; 2000US-0251988P.
05-DEC-2000; 2000US-0256719P.
06-DEC-2000; 2000US-0251479P.
08-DEC-2000; 2000US-0251856P.
08-DEC-2000; 2000US-0251868P.
08-DEC-2000; 2000US-0251869P.
08-DEC-2000; 2000US-0251989P.
08-DEC-2000; 2000US-0251990P.
11-DEC-2000; 2000US-0254097P.
05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-465566/50.
DR N-PSDB; AAS41338.
XX
XX Novel polypeptides and polynucleotides useful for diagnosing, preventing,
PT treating neural, immune system, muscular, reproductive, pulmonary,
PT cardiovascular, renal, proliferative disorders and cancerous diseases.
XX
XX Claim 11; SEQ ID NO 1464; 1180pp; English.
XX
XX The present invention relates to the isolation of novel human enzyme
CC polypeptides, and the cDNA (AAS40785-AAS41684) and genomic sequences
CC encoding them. The enzyme polypeptides of the invention may comprise the
CC functional classes of oxidoreductases, transferases, hydrolases, lyases,
CC isomerases or ligases. The sequences of the invention are useful in the
CC diagnosis, treatment, prevention and/or prognosis of a wide range of
CC disorders including hyperproliferative disorders (e.g. cancer),
CC immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g.
CC arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic
CC disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma),
CC cardiovascular disorders (e.g. atherosclerosis), blood-related disorders
CC (e.g. haemophilia), reproductive disorders (e.g. infertility) and
CC infectious disorders (e.g. influenza). The polynucleotides of the
CC invention can also be used in gene therapy. AAU22915-AAU23814 represent
CC the novel human enzyme polypeptides of the invention. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 149 AA;
Query Match 26.0%; Score 56.5; DB 4; Length 149;
Best Local Similarity 35.3%; Pred. No. 2.9;
Matches 18; Conservative 7; Mismatches 11; Indels 15; Gaps 3;
QY 5 LRLLNKK-----NALRKAAHTSMVNRFRYG-----KPVQSQVOL-KPR 40
Db 52 LRMTLKESAMIAKSRKPHATMIRNKAHYGLHAGWSLWLPVESALQSHQPR 102
RESULT 17
ABB57741
ID ABB57741 standard; protein; 866 AA.
XX
XX ABB57741;
AC
XX
XX 26-MAR-2002 (first entry)
DT
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 15.
DE
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
XX Drosophila melanogaster.
OS
XX
XX WO200171042-A2.
PN
XX
XX 27-SEP-2001.
PD
XX
XX 23-MAR-2001; 2001WO-US009231.
PF
XX
XX 23-MAR-2000; 2000US-0191637P.
PR
XX
XX 11-JUL-2000; 2000US-00614150.
PR
XX
XX (PEKE ) PE CORP NY.
PA
XX
XX Venter JC, Adams M, Li PWD, Myers BW;
PI
XX
XX WPI; 2001-656860/75.
DR
XX
XX N-PSDB; ABL01844.
DR
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
XX
XX Disclosure; SEQ ID NO 15; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 866 AA;
Query Match 24.9%; Score 54; DB 4; Length 866;
Best Local Similarity 33.3%; Pred. No. 61;
Matches 15; Conservative 11; Mismatches 7; Indels 12; Gaps 3;
QY 2 LSNLRILLNKAALRKAAHTSMVNRFRYGK----PVQSQVOLKPRD 41
Db 613 LTNLHILI-----SHTDLMEDFHRFGRLNLYEPVAYRLGWEPD 650
RESULT 18
ABB59259
ID ABB59259 standard; protein; 755 AA.
XX
XX ABB59259;
AC
XX
XX 26-MAR-2002 (first entry)
DT
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 4569.
DE
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
```



XX OS Drosophila melanogaster.  
XX PN WO200171042-A2.  
XX PD 27-SEP-2001.  
XX PF 23-MAR-2001; 2001WO-US009231.  
XX PR 23-MAR-2000; 2000US-0191637P.  
XX PR 11-JUL-2000; 2000US-00614150.  
XX PA (PEKE ) PE CORP NV.  
XX PI Venter JC, Adams M, Li PWD, Myers EW;  
XX DR WPI; 2001-656860/75.  
XX DR N-PSDB; ABL03362.  
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more  
XX PT genes from Drosophila and for elucidating cell signaling and cell-cell  
XX PT interactions.  
XX PS Disclosure; SEQ ID NO 4569; 21pp + Sequence Listing; English.  
XX CS The invention relates to an isolated nucleic acid detection reagent  
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is  
XX CC useful in developmental biology and in elucidating cell signalling and  
XX CC cell-cell interactions in higher eukaryotes for the development of  
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention  
XX CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
XX CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-  
XX CC ABB72072). The sequence data for this patent did not form part of the  
XX CC printed specification, but was obtained in electronic format directly  
XX CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX CS Sequence 755 AA;  
XX SQ

Query Match 24.7%; Score 53.5; DB 4; Length 755;  
Best Local Similarity 42.5%; Pred. No. 62;  
Matches 17; Conservative 5; Mismatches 15; Indels 3; Gaps 2;  
QY 2 LSNRLILNKAALRKAKHTSMVRNFRYKPKVQSQVQLKPRD 41  
DB 11 LSDLRRELE-VLRKAHFELDLHLYFTGTG-QPEAEAKPRD 47

RESULT 19  
ABO84066  
ID ABO84066 standard; protein; 457 AA.  
AC ABO84066;  
XX 29-JUL-2004 (first entry)  
XX DE Pseudomonas aeruginosa polypeptide #16241.  
XX KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.  
XX OS Pseudomonas aeruginosa.  
XX PN US6551795-B1.  
XX PD 22-APR-2003.  
XX PF 18-FEB-1999; 99US-00252991.  
XX PR 18-FEB-1998; 98US-0074788P.  
XX PR 27-JUL-1998; 98US-0094190P.  
XX PA (GENO-) GENOME THERAPEUTICS CORP.  
XX PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;

XX DR WPI; 2003-615309/58.  
XX DR N-PSDB; ABD17637.  
XX PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,  
XX PT useful as molecular targets for diagnostics, prophylaxis and treatment of  
XX PT pathological conditions resulting from bacterial infection.  
XX PS Disclosure; SEQ ID NO 32812; 455pp; English.  
XX CC The invention relates to Pseudomonas aeruginosa polypeptides and the  
XX CC polynucleotides encoding them. The sequences are useful in diagnosis and  
XX CC therapy of pathological conditions, as molecular targets for diagnostics,  
XX CC prophylaxis and treatment of pathological conditions resulting from a  
XX CC bacterial infection, for evaluating a compound, such as a polypeptide,  
XX CC for the ability to bind a P. aeruginosa nucleic acid, as components of  
XX CC effective antibacterial targets, as targets for antibacterial drugs,  
XX CC including anti-P. aeruginosa drugs, as templates for recombinant  
XX CC production of P. aeruginosa-derived peptides or polypeptides, as target  
XX CC components for diagnosis and/or treatment of P. aeruginosa-caused  
XX CC infection, and in detection of P. aeruginosa sequences or other sequences  
XX CC of Pseudomonas species using biochip technology. Sequences ABO67826-  
XX CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The  
XX CC sequence data for this patent did not form part of the printed  
XX CC specification but was obtained in electronic format from USPTO at  
XX CC seqdata.uspto.gov/sequence.html  
XX CS Sequence 457 AA;  
XX SQ

Query Match 24.4%; Score 53; DB 7; Length 457;  
Best Local Similarity 34.4%; Pred. No. 40;  
Matches 11; Conservative 6; Mismatches 15; Indels 0; Gaps 0;  
QY 7 ILLNKAALRKAKHTSMVRNFRYKPKVQSQVQLK 38  
DB 176 VALDAAVRRARLVRFRGLGRPAQLYVETR 207

RESULT 20  
ABU50395  
ID ABU50395 standard; protein; 593 AA.  
XX AC ABU50395;  
XX DT 19-JUN-2003 (first entry)  
XX DE Protein encoded by Prokaryotic essential gene #35922.  
XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.  
XX OS Yersinia pestis.  
XX PN WO200277183-A2.  
XX PD 03-OCT-2002.  
XX PF 21-MAR-2002; 2002WO-US009107.  
XX PR 21-MAR-2001; 2001US-00815242.  
XX PR 06-SEP-2001; 2001US-00948993.  
XX PR 25-OCT-2001; 2001US-0342923P.  
XX PR 08-FEB-2002; 2002US-00072851.  
XX PR 06-MAR-2002; 2002US-0362699P.  
XX PA (ELIT-) ELITRA PHARM INC.  
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zykkind JW;  
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
XX DR WPI; 2003-029926/02.  
XX DR N-PSDB; ACA54265.  
XX PT New antisense nucleic acids, useful for identifying proteins or screening



CC effective antibacterial targets, as targets for antibacterial drugs,  
CC including anti-P. aeruginosa drugs, as templates for recombinant  
CC production of P. aeruginosa-derived peptides or polypeptides, as target  
CC components for diagnosis and/or treatment of P. aeruginosa-caused  
CC infection, and in detection of P. aeruginosa sequences or other sequences  
CC of Pseudomonas species using biochip technology. Sequences ABO67826-  
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format from USPTO at  
CC seqdata.uspto.gov/sequence.html  
XX  
XX Sequence 218 AA;  
SQ

Query Match 24.0%; Score 52; DB 7; Length 218;  
Best Local Similarity 33.3%; Pred. No. 23;  
Matches 16; Conservative 6; Mismatches 20; Indels 6; Gaps 1;  
QY 1 MLSNRLILNKALRK-----AHTSMVRNFRYKPVQSQVOLKPRDL 42  
DB 116 MDDLPKATRHAGRRPFISRSIAHLLVEGEAEHGKPLASLEQLTPREV 163

RESULT 23  
ABG29304  
ID ABG29304 standard; protein; 774 AA.  
XX  
AC ABG29304;  
XX  
DT 13-FEB-2002 (first entry)  
XX  
DE Novel human diagnostic protein #29295.  
XX  
DE Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.  
XX  
OS Homo sapiens.  
XX  
PN WO200175067-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 30-MAR-2001; 2001WO-US008631.  
XX  
PR 31-MAR-2000; 2000US-00540217.  
XX  
PR 23-AUG-2000; 2000US-00649167.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Drmanac RT, Liu C, Tang YT;  
XX  
XX WPI; 2001-639362/73.  
DR N-PSDB; AAS93491.  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX  
XX Claim 20; SEQ ID NO 59663; 103pp; English.  
PS  
XX  
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits and to assess  
CC biodiversity.  
XX  
XX Claim 20; SEQ ID NO 59663; 103pp; English.  
PS  
XX  
XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (AB101840-ABL16175) and the encoded proteins (ABB57737-  
CC ABB72072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 1124 AA;  
SQ

Query Match 24.0%; Score 52; DB 4; Length 774;  
Best Local Similarity 31.0%; Pred. No. 11e+02;  
Matches 13; Conservative 9; Mismatches 18; Indels 2; Gaps 1;  
QY 1 MLSNRLILNKALRKRAHTSMVRNF--RYKPVQSQVOLKPR 40  
DB 624 LTTFRQLQTEPLRPALEASCEEYSAKPGFVFKLDYQLPPR 665

RESULT 24  
ABB59241  
ID ABB59241 standard; protein; 1124 AA.  
XX  
AC ABB59241;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster polypeptide SEQ ID NO 4515.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US009231.  
XX  
PR 23-MAR-2000; 2000US-0191637P.  
XX  
PR 11-JUL-2000; 2000US-00614150.  
XX  
XX (PEKE ) PE CORP NY.  
XX  
XX Venter JC, Adams M, Li PWD, Myers EW;  
PI  
PI WPI; 2001-656860/75.  
DR N-PSDB; ABL03344.  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions.  
XX  
XX Disclosure; SEQ ID NO 4515; 21pp + Sequence Listing; English.  
PS  
XX  
XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (AB101840-ABL16175) and the encoded proteins (ABB57737-  
CC ABB72072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 1124 AA;  
SQ

Query Match 24.0%; Score 52; DB 4; Length 1124;  
Best Local Similarity 39.1%; Pred. No. 1.7e+02;  
Matches 9; Conservative 5; Mismatches 9; Indels 0; Gaps 0;





ABR53783  
ID ABR53783 standard; protein; 705 AA.  
XX  
AC ABR53783;  
XX  
DT 20-JUN-2003 (first entry)  
XX  
DE Protein sequence #SEQ ID 2431.  
XX  
KW Multiprotein complex; eukaryote; drug target; diagnosis.  
XX  
OS Saccharomyces cerevisiae.  
XX  
PN EP1258494-A1.  
XX  
PD 20-NOV-2002.  
XX  
PF 20-DEC-2001; 2001EP-00130253.  
XX  
PR 15-MAY-2001; 2001EP-00111774.  
XX  
PA (CELL-) CELLZONE AG.  
XX  
PI Bauer A, Gavin A, Grandi P, Krause R, Kruse UD, Kuester BD;  
PI Marzloch M, Schultz JD, Superti-Furga G;  
DR WPI; 2003-250078/25.  
DR N-PSDB; ACC61825.  
XX  
PT New isolated protein complexes useful for diagnosing a disease or  
PT disorder, or as a target for an active agent of a pharmaceutical,  
PT preferably a drug target in the treatment or prevention of disease or  
PT disorder.  
XX  
PS Disclosure; SEQ ID NO 2431; 17pp + Sequence Listing; English.  
XX  
CC The invention relates to multiprotein complexes from eukaryotes. Proteins  
CC of the invention and DNA sequences encoding them are given in records  
CC ABR53783 and ACC61825 respectively. The complexes are  
CC obtainable by using a protein as a bait and isolating the set of proteins  
CC which is attached thereto from cells. Such protein complexes may comprise  
CC up to 30 distinct proteins. Protein complexes of the invention are useful  
CC for diagnosing a disease or disorder, or as a target for an active agent  
CC of a pharmaceutical, preferably a drug target in the treatment or  
CC prevention of a disease or disorder. Note: The sequence data for this  
CC patent is not represented in the printed specification, but is based on  
CC sequence information supplied by the European Patent Office. The complete  
CC document is available on CD-ROM  
XX  
SQ Sequence 705 AA;  
Query Match 23.5%; Score 51; DB 6; Length 705;  
Best Local Similarity 36.8%; Pred. No. 1.4e+02;  
Matches 14; Conservative 5; Mismatches 19; Indels 0; Gaps 0;  
OY 5 LRLLNKAALRKHAHTSMVRNFRYKPVQSQVQLKPRDL 42  
Db 643 LRTAKSFELLRKAQASMSVKFGFKPLRDDAFLESRL 680  
A pharmaceutical, preferably a drug target in the treatment or prevention of a disease or disorder. Note: The sequence data for this patent is not represented in the printed specification, but is based on sequence information supplied by the European Patent Office. The complete document is available on CD-ROM  
RESULT 30  
ADK63830  
ID ADK63830 standard; protein; 705 AA.  
XX  
AC ADK63830;  
XX  
DT 06-MAY-2004 (first entry)  
XX  
DE Disease treating protein complex-derived protein #1460.  
DE  
XX  
KW protein complex; drug target; diagnosis.  
XX  
OS Unidentified.

XX  
PN EP1338608-A2.  
XX  
PD 27-AUG-2003.  
XX  
PF 20-DEC-2002; 2002EP-00102902.  
XX  
PR 20-DEC-2001; 2001EP-00130253.  
XX  
PA (CELL-) CELLZONE AG.  
XX  
PI Bauer A, Gavin A, Superti-Furga G, Kuester B, Schultz J;  
PI Marzloch M, Grandi P, Krause R, Kruse U, Merino A, Bauch A;  
PI Michon A, Leutwein C, Rick J;  
XX  
DR WPI; 2003-638460/61.  
DR N-PSDB; ADK63831.  
XX  
PT New proteins and protein complexes from eukaryotes, useful as targets in  
PT drug screening, or in diagnosing or screening for the presence of a  
PT disease or disorder, or a predisposition for developing a disease or  
PT disorder in a subject.  
XX  
PS Disclosure; SEQ ID NO 2919; 13pp; English.  
XX  
CC The invention relates to novel protein complexes comprising a first and a  
CC second protein, or its derivative, fragment, homologue or variant. The  
CC proteins are selected from given protein complexes, which are not defined  
CC in the specification. The variants are encoded by nucleic acids that  
CC hybridize to the nucleic acids encoding the proteins under low stringency  
CC conditions. The protein complexes are useful as targets for an active  
CC agent of a pharmaceutical. These protein complexes are particularly  
CC useful as drug targets for the treatment or preventing of a disease or  
CC disorder. The complexes and methods above are useful in diagnosing or  
CC screening for the presence of a disease or disorder or a predisposition  
CC for developing a disease or disorder in a subject. These are also useful  
CC in screening for a drug for treatment or prevention of a disease or  
CC disorder. The molecule that modulates the amount, activity or protein  
CC components of the complex is useful for the manufacture of a medicament  
CC for the treatment or prevention of a disease or disorder. This sequence  
CC corresponds to a protein of the invention. (Note: the sequence data for  
CC this patent did not form part of the printed specification but was  
CC obtained from the EPO in electronic format).  
XX  
SQ Sequence 705 AA;  
Query Match 23.5%; Score 51; DB 7; Length 705;  
Best Local Similarity 36.8%; Pred. No. 1.4e+02;  
Matches 14; Conservative 5; Mismatches 19; Indels 0; Gaps 0;  
OY 5 LRLLNKAALRKHAHTSMVRNFRYKPVQSQVQLKPRDL 42  
Db 643 LRTAKSFELLRKAQASMSVKFGFKPLRDDAFLESRL 680  
A pharmaceutical, preferably a drug target in the treatment or prevention of a disease or disorder. Note: The sequence data for this patent is not represented in the printed specification, but is based on sequence information supplied by the European Patent Office. The complete document is available on CD-ROM  
RESULT 31  
AAR96126  
ID AAR96126 standard; protein; 1169 AA.  
XX  
AC AAR96126;  
XX  
DT 16-OCT-2003 (revised)  
DT 28-OCT-1996 (first entry)  
XX  
DE Bacillus thuringiensis crystal protein.  
XX  
KW Bacillus thuringiensis; crystal protein; delta toxin; insecticide;  
KW Lepidoptera; Coleoptera; crop protection; rice; wheat; beans; tea;  
KW sugarcane; cauliflower; cabbage; apple; citrus fruit.  
XX  
OS Bacillus thuringiensis; var. japonensis.  
OS strain N141.  
XX

PN EP711834-A2.  
XX  
PD 15-MAY-1996.  
XX  
PF 13-OCT-1995; 95EP-00307293.  
XX  
PR 14-OCT-1994; 94JP-00276082.  
XX  
PA (NISC) NISSAN CHEM IND LTD.  
XX  
XX Iizuka T, Tagawa M, Arai S, Niizeki M, Miyake T;  
XX WPI: 1996-232099/24.  
DR N-PSDB; AAT27148.  
XX  
XX B. thuringiensis var japonensis strain N141 insecticidal crystal protein  
PT - used to protect plant from damage by pest, partic, lepidoptera or  
PT coleoptera insects.  
XX  
XX Claim 3; Page 12-16; 20pp; English.  
XX  
XX The crystal protein is toxic to insects of the order Lepidoptera or  
CC Coleoptera, it is therefore useful in the protection of crops which are  
CC subject to infestation with lepidopterous or coleopterous insect pests.  
CC  
CC The crystal protein is formulated into a spray so that the protein agent  
CC is applied to crops at an amount of 0.1 to 5kg per hectare. Plants to be  
CC protected by such methods include vegetables such as cauliflower and  
CC cabbage, fruit trees such as citrus and apples, grains such as rice,  
CC wheat and beans and industrial crops such as tea and sugarcane. (Updated  
CC on 16-OCT-2003 to standardise OS field)  
XX  
XX Sequence 1169 AA;  
SQ  
Query Match 23.5%; Score 51; DB 2; Length 1169;  
Best Local Similarity 25.5%; Pred. No. 2.6e+02;  
Matches 14; Conservative 12; Mismatches 15; Indels 14; Gaps 2;  
QY 2 LSNLRLILN-----KAALRKHAHTSMVRNFRY---KPVQSQVQLKPRDL 42  
DB 359 LSNLQILVNYQTNGAWGRVRYHYLHSSIIQEKSYGLLSDPVGANINVQNNDI 413  
RESULT 32  
AAAY75466  
ID AAAY75466 standard; protein; 138 AA.  
XX  
AC AAAY75466;  
XX  
DT 21-MAR-2000 (first entry)  
XX  
DE Neisseria meningitidis ORF 715 protein sequence SEQ ID NO:2406.  
XX  
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;  
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;  
KW antibacterial; gene therapy.  
XX  
OS Neisseria meningitidis.  
XX  
PN WO9957280-A2.  
XX  
PD 11-NOV-1999.  
XX  
PF 30-APR-1999; 99WO-US009346.  
XX  
PR 01-MAY-1998; 98US-0083758P.  
PR 31-JUL-1998; 98US-0094869P.  
PR 02-SEP-1998; 98US-0098994P.  
PR 02-SEP-1998; 98US-0099062P.  
PR 09-OCT-1998; 98US-0103749P.  
PR 09-OCT-1998; 98US-0103794P.  
PR 09-OCT-1998; 98US-0103796P.  
PR 25-FEB-1999; 99US-0121528P.  
XX

PA (CHIR) CHIRON CORP.  
PA (GENO-) INST GENOMIC RES.  
XX  
PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;  
PI Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;  
PI Tettelin H, Venter JC;  
XX  
DR WPI: 2000-062150/05.  
DR N-PSDB; AAZ54228.  
XX  
XX Novel Neisserial polypeptides predicted to be useful antigens for  
PT vaccines and diagnostics.  
XX  
XX Claim 2; Page 1154; 1453pp; English.  
XX  
XX AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAZ74253 to AAZ75941  
CC represent novel Neisseria meningitis and N. gonorrhoeae polynucleotides  
CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent  
CC PCR primers used in the exemplification of the present invention. The  
CC polypeptides, the polynucleotides, antibodies and compositions of the  
CC invention can be used as vaccines, as diagnostic reagents, and as  
CC immunogenic compositions. The polypeptides can be used in the manufacture  
CC of medicaments for treating or preventing infection due to Neisserial  
CC bacteria (e.g. meningitis and septicemia), to detect the presence of  
CC Neisseria bacteria, or to raise antibodies. They may also be used to  
CC screen for agonists or antagonists, which may themselves have use as  
CC antibacterial agents. The polynucleotides of the invention may also be  
CC used in gene therapy protocols  
XX  
XX Sequence 138 AA;  
SQ  
Query Match 23.3%; Score 50.5; DB 3; Length 138;  
Best Local Similarity 30.4%; Pred. No. 22;  
Matches 14; Conservative 5; Mismatches 14; Indels 13; Gaps 1;  
QY 6 RILLNKAALRKHAHTSMVRNFRY-----GKPVSQVQLK 38  
DB 26 RYLLMRLSETMHTAVKLNFRYAGRPKWGLKLYRGKPLSDSGRLX 71  
RESULT 33  
ABB64170  
ID ABB64170 standard; protein; 296 AA.  
XX  
AC ABB64170;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster polypeptide SEQ ID NO 19302.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US009231.  
XX  
PR 23-MAR-2000; 2000US-0191637P.  
PR 11-JUL-2000; 2000US-00614150.  
XX  
XX (PEKS) PE CORP NY.  
XX  
XX Venter JC, Adams M, Li PWD, Myers EW;  
PI WPI: 2001-656860/75.  
DR N-PSDB; ABL08273.  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signaling and cell-cell

PT interactions.

PS Disclosure; SEQ ID NO 19302; 21pp + Sequence Listing; English.

XX

XX The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from *Drosophila*. The invention is

CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-

CC ABB572072). The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly

CC from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences)

XX

SQ Sequence 296 AA;

Query Match 23.3%; Score 50.5; DB 4; Length 296;

Best Local Similarity 44.4%; Pred. No. 57;

Matches 12; Conservative 5; Mismatches 9; Indels 1; Gaps 1;

QY 2 LSNLRILLNKAALRKAH-TSMVNFERY 27

Db 76 LEFISILFKKAHLKSHRLRKVKNFY 102

RESULT 34

ABM67661

ID ABM67661 standard; protein; 533 AA.

XX

AC ABM67661;

XX

XX 20-NOV-2003 (first entry)

DT

XX

DE Photarhabdus luminescens protein sequence #759.

XX

XX Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;

KW detection; food; gene expression; plant; animal; microorganism; toxin;

KW antibiotic; biopesticide; virulence factor; disease model; plague;

KW whooping cough.

XX

XX Photarhabdus luminescens.

OS

XX WO200294867-A2.

PN

XX 28-NOV-2002.

PD

XX

XX 07-FEB-2002; 2002WO-IB003040.

PF

XX

XX 07-FEB-2001; 2001FR-00001659.

PR

XX

XX (INSP ) INST PASTEUR.

PA (CNRS ) CNRS CENT NAT RECH SCI.

PA

XX Duchaud E, Taourit S, Glaser P, Frangeul L, Kunat F, Danchin A;

PI Buchrieser C;

PI

XX WPI; 2003-148459/14.

DR

XX

XX Genomic sequence of *Photarhabdus luminescens* and encoded polypeptides,

PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.

PT

XX

PS Claim 2; SEQ ID NO 758; 1205pp; French.

PS

XX

XX The invention relates to the isolation of genes and their encoded

CC proteins from *Photarhabdus luminescens*. The isolated sequences are

CC sources of probes and primers for detecting the genome of *P. luminescens*

CC and related species; to study polymorphisms; for gene analysis and for

CC detection/amplification of the genes. Antibodies (Ab) raised against the

CC polypeptides encoded by the genes are used for detection/identification

CC of *P. luminescens*, e.g. in foods. The genes, proteins, Ab and cells that

CC carry a gene-containing vector are used to select compounds that

CC modulate, regulate, induce or inhibit expression of the genes in plants,

CC

CC animals or microorganisms other than *P. luminescens* and are able to alter

CC response or sensitivity to toxins and antibiotics produced by *P.*

CC *luminescens*. Cells transformed to express the genes are useful for

CC recombinant production of the proteins, particularly toxins and

CC antibacterials useful as insecticides, bactericides and fungicides. The

CC genes, proteins, vectors containing the genes and Ab are also useful

CC therapeutically (to treat microbial infection by bacteria or fungi that

CC are sensitive to *P. luminescens*-encoded toxins or antibiotics) and as

CC biopesticides. Other uses of the genes and the proteins are as virulence

CC factors and for identifying targets of human diseases for which *P.*

CC *luminescens* is a model (particularly plague and whooping cough). This

CC sequence represents one of the isolated *P. luminescens* proteins

XX

SQ Sequence 533 AA;

Query Match 23.3%; Score 50.5; DB 6; Length 533;

Best Local Similarity 47.8%; Pred. No. 1.2e+02;

Matches 11; Conservative 6; Mismatches 5; Indels 1; Gaps 1;

QY 2 LSNLRILLNKA-ALRKAHTSMVR 23

Db 285 LSSIRLVSTGMALRKQHVSMIK 307

RESULT 35

ABP62760

ID ABP62760 standard; protein; 2379 AA.

XX

AC ABP62760;

XX

XX 23-OCT-2002 (first entry)

DT

XX

DE *S. roseosporus* daptomycin non-ribosomal peptide synthetase DptD.

XX

XX Daptomycin biosynthetic gene cluster; thioesterase; antibacterial;

KW fungicide; viricide; antiparasitic; immunomodulator; antilipemic;

KW cytostatic; gene therapy; antimitotic; immunomodulatory; siderophore;

KW anti-cholesterolemic; agrochemical; non-ribosomal peptide synthetase;

KW NRPS; DptD.

XX

XX Streptomyces roseosporus.

OS

XX WO200259322-A2.

PN

XX 01-AUG-2002.

PD

XX

XX 17-OCT-2001; 2001WO-US032354.

PF

XX

XX 17-OCT-2000; 2000US-0240879P.

PR

XX 28-FEB-2001; 2001US-0272207P.

PR

XX 06-AUG-2001; 2001US-0310385P.

PR

XX

XX (MIAO/) MIAO V P W.

PA (ERIA/) BRIAN P.

PA (BALT/) BALTZ R H.

PA (SILV/) SILVA C J.

PA

XX Miao VPW, Brian P, Baltz RH, Silva CJ;

PI

XX WPI; 2002-599794/64.

DR

XX

XX Isolated nucleic acid molecule from a bacterial daptomycin biosynthetic

PT gene cluster encoding a thioesterase or thioesterase domain, useful for

PT generating novel linear and cyclic peptides, and products in a cell.

PT

XX

XX Claim 7; Page 165-166; 227pp; English.

PS

XX

XX The invention relates to a novel isolated nucleic acid molecule

CC comprising a sequence that encodes a thioesterase or thioesterase domain,

CC derived from a bacterial daptomycin biosynthetic gene cluster. The

CC proteins of the invention have antibacterial, fungicide, viricide,

CC antiparasitic, immunomodulator, antilipemic, and cytostatic activity. The

CC polynucleotides may have a use in gene therapy. The compositions and

CC



CC methods of the present invention are useful for generating novel linear  
CC an daptomycin non-ribosomal peptide synthetase (NRPS) to be used as new  
CC compounds or in producing new compounds, such as antibiotics,  
CC antifungals, antivirals, antiparasitics, antimitotics, antitumour agents,  
CC immunomodulatory agents, anti-cholesterolemic agents, siderophores,  
CC agrochemicals and cytostatics. The sequence represents a S. roseosporus  
CC daptomycin non-ribosomal peptide synthetase of the invention  
XX  
SQ Sequence 2379 AA;  
Query Match 23.3%; Score 50.5; DB 5; Length 2379;  
Best Local Similarity 38.7%; Pred. No. 7.4e+02;  
Matches 12; Conservative 8; Mismatches 10; Indels 1; Gaps 1;  
QY 12 AALRKAHTSMVRNFRYKPKVQSQVQLKPRDL 42  
DB 55 AALLRRHANLRAAFRYER-LQRPVQIIPREV 84  
RESULT 36  
ADJ72172  
ID ADJ72172 standard; protein; 2379 AA.  
AC ADJ72172;  
XX  
DT 06-MAY-2004 (first entry)  
DE Streptomyces roseosporus DptD protein.  
XX  
DE antibacterial; gene therapy; daptomycin biosynthesis gene cluster;  
KW daptomycin non-ribosomal peptide synthetase; DptBC;  
KW gram-positive bacterial infection.  
XX  
OS Streptomyces roseosporus.  
XX  
PN WO2003014297-A2.  
XX  
PD 20-FEB-2003.  
XX  
PF 31-JUL-2002; 2002WO-US024310.  
XX  
PR 06-AUG-2001; 2001US-0310385P.  
PR 17-OCT-2001; 2001WO-US032354.  
PR 10-MAY-2002; 2002US-0379866P.  
XX  
PA (CUBI-) CUBIST PHARM INC.  
XX  
PI Miao VPW, Brian P, Baltz RH, Coeffet-Legal MF;  
XX  
DR WPI; 2003-268192/26.  
DR N-PSDB; ADJ72363.  
XX  
PT New isolated nucleic acid molecule encoding a daptomycin non-ribosomal  
PT peptide synthetase, useful for treatment of a gram-positive bacterial  
PT infection of skeletal muscle, skin, bloodstream, kidneys, heart, lung and  
PT bone.  
XX  
PS Disclosure; SEQ ID NO 7; 292pp; English.  
XX  
SS The invention relates to new isolated nucleic acid (NA) molecules from  
CC the Streptomyces roseosporus daptomycin biosynthesis gene cluster,  
CC especially a daptomycin non-ribosomal peptide synthetase (NRPS) or its  
CC subunit, where the (NA) molecule encodes DptBC, and is not pRB159. The  
CC methods and compositions of the present invention are useful for  
CC treatment of a gram-positive bacterial infection of any organ or tissue  
CC in the body, including skeletal muscle, skin, bloodstream, kidneys,  
CC heart, lung and bone. This sequence represents the daptomycin  
CC biosynthesis protein DptD.  
XX  
SQ Sequence 2379 AA;  
Query Match 23.3%; Score 50.5; DB 7; Length 2379;  
Best Local Similarity 38.7%; Pred. No. 7.4e+02;  
Matches 12; Conservative 8; Mismatches 10; Indels 1; Gaps 1;  
QY 12 AALRKAHTSMVRNFRYKPKVQSQVQLKPRDL 42  
DB 55 AALLRRHANLRAAFRYER-LQRPVQIIPREV 84  
RESULT 37  
ABP73265  
ID ABP73265 standard; protein; 564 AA.  
XX  
AC ABP73265;  
XX  
DT 30-JAN-2003 (first entry)  
DE Candida albicans essential protein SEQ ID NO 7102.  
XX  
DE Fungus; yeast; tetracycline; promoter; GRACE strain; biosynthesis;  
KW signal transduction; DNA replication; cell division; growth;  
KW proliferation; Candida albicans; fungicide; antifungal.  
XX  
OS Candida albicans.  
XX  
PN WO200253728-A2.  
XX  
PD 11-JUL-2002.  
XX  
PF 26-DEC-2001; 2001WO-US049486.  
XX  
PR 29-DEC-2000; 2000US-0259128P.  
PR 20-FEB-2001; 2001US-00792024.  
PR 22-AUG-2001; 2001US-0314050P.  
XX  
XX (ELIT-) ELITRA PHARM INC.  
XX  
XX Roemer T, Jiang B, Boone C, Bussey H, Ohlsen KL;  
XX  
DR WPI; 2002-566694/60.  
DR N-PSDB; ABZ31815.  
XX  
XX Constructing strains for identifying gene products as effective targets  
PT for therapeutic intervention, by inactivating in the strain one allele of  
PT a gene and placing other allele of the gene under conditional expression.  
XX  
PS Claim 44; SEQ ID NO 7102; 167pp + Sequence Listing; English.  
XX  
CC The invention relates to constructing (M1) a strain of diploid fungal  
CC cells in which both alleles of a gene are modified, comprising modifying  
CC one allele by insertion or replacement by a cassette having an  
CC expressible selectable marker and modifying other allele by  
CC recombination, of a promoter replacement fragment with a heterologous  
CC promoter, so that expression of the second allele is regulated by the  
CC promoter. (M1) is useful for constructing a strain of diploid fungal  
CC cells in which both alleles of a gene are modified. The diploid fungal  
CC cells having both alleles modified are useful for identifying a gene that  
CC is essential to the survival or growth of a fungus, a gene that  
CC that contributes to the virulence and/or pathogenicity of a fungus, a gene  
CC agent, an antifungal agent that inhibits the growth of a diploid fungus  
CC and for identifying a therapeutic agent for treatment of a mammalian  
CC disease. (M1) is useful for identifying a compound which modulates the  
CC activity of a gene product, preferably enzymatic activity, carbon  
CC compound catabolism, biosynthetic, transporter, transcriptional,  
CC translational, signal transduction, DNA replication and cell division  
CC activity. The method is useful for identifying a compound having the  
CC ability to inhibit growth or proliferation of C. albicans cells and for  
CC treating infection by C. albicans. The present sequence is that of an  
CC essential Candida albicans protein used in the method of the invention.  
CC Note: The sequence data for this patent is not represented in the printed  
CC specification but is based on sequence information supplied to Derwent by  
CC the European Patent Office  
XX

Best Local Similarity 38.7%; Pred. No. 7.4e+02;  
Matches 12; Conservative 8; Mismatches 10; Indels 1; Gaps 1;  
QY 12 AALRKAHTSMVRNFRYKPKVQSQVQLKPRDL 42  
DB 55 AALLRRHANLRAAFRYER-LQRPVQIIPREV 84  
RESULT 37  
ABP73265  
ID ABP73265 standard; protein; 564 AA.  
XX  
AC ABP73265;  
XX  
DT 30-JAN-2003 (first entry)  
DE Candida albicans essential protein SEQ ID NO 7102.  
XX  
DE Fungus; yeast; tetracycline; promoter; GRACE strain; biosynthesis;  
KW signal transduction; DNA replication; cell division; growth;  
KW proliferation; Candida albicans; fungicide; antifungal.  
XX  
OS Candida albicans.  
XX  
PN WO200253728-A2.  
XX  
PD 11-JUL-2002.  
XX  
PF 26-DEC-2001; 2001WO-US049486.  
XX  
PR 29-DEC-2000; 2000US-0259128P.  
PR 20-FEB-2001; 2001US-00792024.  
PR 22-AUG-2001; 2001US-0314050P.  
XX  
XX (ELIT-) ELITRA PHARM INC.  
XX  
XX Roemer T, Jiang B, Boone C, Bussey H, Ohlsen KL;  
XX  
DR WPI; 2002-566694/60.  
DR N-PSDB; ABZ31815.  
XX  
XX Constructing strains for identifying gene products as effective targets  
PT for therapeutic intervention, by inactivating in the strain one allele of  
PT a gene and placing other allele of the gene under conditional expression.  
XX  
PS Claim 44; SEQ ID NO 7102; 167pp + Sequence Listing; English.  
XX  
CC The invention relates to constructing (M1) a strain of diploid fungal  
CC cells in which both alleles of a gene are modified, comprising modifying  
CC one allele by insertion or replacement by a cassette having an  
CC expressible selectable marker and modifying other allele by  
CC recombination, of a promoter replacement fragment with a heterologous  
CC promoter, so that expression of the second allele is regulated by the  
CC promoter. (M1) is useful for constructing a strain of diploid fungal  
CC cells in which both alleles of a gene are modified. The diploid fungal  
CC cells having both alleles modified are useful for identifying a gene that  
CC is essential to the survival or growth of a fungus, a gene that  
CC that contributes to the virulence and/or pathogenicity of a fungus, a gene  
CC agent, an antifungal agent that inhibits the growth of a diploid fungus  
CC and for identifying a therapeutic agent for treatment of a mammalian  
CC disease. (M1) is useful for identifying a compound which modulates the  
CC activity of a gene product, preferably enzymatic activity, carbon  
CC compound catabolism, biosynthetic, transporter, transcriptional,  
CC translational, signal transduction, DNA replication and cell division  
CC activity. The method is useful for identifying a compound having the  
CC ability to inhibit growth or proliferation of C. albicans cells and for  
CC treating infection by C. albicans. The present sequence is that of an  
CC essential Candida albicans protein used in the method of the invention.  
CC Note: The sequence data for this patent is not represented in the printed  
CC specification but is based on sequence information supplied to Derwent by  
CC the European Patent Office  
XX

XX SQ Sequence 746 AA;  
Query Match 23.0%; Score 50; DB 4; Length 746;  
Best Local Similarity 34.9%; Pred. No. 2.1e+02;  
Matches 15; Conservative 5; Mismatches 21; Indels 2; Gaps 1;

QY 1 MLSNRILLINKAALRKAHTSMVRNFRYK--PVSOVOLKPRD 41  
| | : | : | : | : | : | : | : | : | : | : | : | : |  
DB 155 MQSQFSVLVNESLQKAHLDEARTLLHGTGTHQHVLEIARD 197

RESULT 39  
ID AAB94042 standard; protein; 975 AA.  
AC AAB94042;  
DT 26-JUN-2001 (first entry)  
DE Human protein sequence SEQ ID NO:14199.  
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.  
OS Homo sapiens.  
PN EP1074617-A2.  
PD 07-FEB-2001.  
PF 28-JUL-2000; 2000EP-00116126.  
PR 29-JUL-1999; 99JP-00248036.  
PR 27-AUG-1999; 99JP-00300253.  
PR 11-JAN-2000; 2000JP-00118776.  
PR 02-MAY-2000; 2000JP-00183767.  
PR 09-JUN-2000; 2000JP-00241899.  
PA (HELI-) HELIX RES INST.  
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
DR WPI; 2001-318749/34.

XX PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.  
XX PS Claim 8; SEQ ID NO 14199; 2537pp + Sequence Listing; English.  
XX CC The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises:  
CC (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the  
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
CC of an oligonucleotide comprising a sequence complementary to the  
CC complementary strand of a polynucleotide which comprises a 5'-end  
CC sequence and an oligonucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in the  
CC specification. The primer sets can be used in antisense therapy and in  
CC gene therapy. The primers are useful for synthesising polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893  
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent  
CC oligonucleotides, all of which are used in the exemplification of the  
CC present invention

XX SQ Sequence 564 AA;  
Query Match 23.0%; Score 50; DB 5; Length 564;  
Best Local Similarity 29.1%; Pred. No. 1.5e+02;  
Matches 16; Conservative 7; Mismatches 17; Indels 16; Gaps 1;

QY 2 LSNRLILANKAALR-----KAHTSMVRNFRYGKPVQSVOVKPR 40  
| | : | : | : | : | : | : | : | : | : | : | : | : |  
DB 424 LSDELLELVIAVRWIEKFELQTINFLAYTYQWKNVTGNAINSTSLDSR 478

RESULT 38  
ID AAB93938 standard; protein; 746 AA.  
AC AAB93938;  
DT 26-JUN-2001 (first entry)  
DE Human protein sequence SEQ ID NO:13945.  
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.  
OS Homo sapiens.  
PN EP1074617-A2.  
PD 07-FEB-2001.  
PF 28-JUL-2000; 2000EP-00116126.  
PR 29-JUL-1999; 99JP-00248036.  
PR 27-AUG-1999; 99JP-00300253.  
PR 11-JAN-2000; 2000JP-00118776.  
PR 02-MAY-2000; 2000JP-00183767.  
PR 09-JUN-2000; 2000JP-00241899.  
PA (HELI-) HELIX RES INST.  
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
DR WPI; 2001-318749/34.

XX PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.  
XX PS Claim 8; SEQ ID NO 13945; 2537pp + Sequence Listing; English.  
XX CC The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises:  
CC (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the  
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
CC of an oligonucleotide comprising a sequence complementary to the  
CC complementary strand of a polynucleotide which comprises a 5'-end  
CC sequence and an oligonucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in the  
CC specification. The primer sets can be used in antisense therapy and in  
CC gene therapy. The primers are useful for synthesising polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893  
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent  
CC oligonucleotides, all of which are used in the exemplification of the  
CC present invention

Search completed: December 18, 2004, 02:45:42  
Job time : 117.179 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 18, 2004, 02:37:41 ; Search time 28.6667 Seconds  
(without alignments)  
99.477 Million cell updates/sec

Title: US-08-765-244-22  
Perfect score: 217  
Sequence: 1 MLSNRLILNKALKAHKAHTS.....NFRYKPVQSVOLKPRDL 43

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/prodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/prodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/prodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/prodata/1/iaa/PCITUS\_COMB.pep.\*  
6: /cgn2\_6/prodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	141	65.0	354	US-09-538-092-833	Sequence 833, Appl
2	92	42.4	32	US-08-373-190-29	Sequence 29, Appl
3	92	42.4	32	US-08-438-190A-29	Sequence 29, Appl
4	92	42.4	32	US-08-350-215-29	Sequence 29, Appl
5	92	42.4	32	US-09-287-145A-29	Sequence 29, Appl
6	92	42.4	32	US-09-556-111-29	Sequence 29, Appl
7	56	25.8	226	US-09-270-767-56794	Sequence 56794, A
8	56	25.8	602	US-09-270-767-41564	Sequence 41564, A
9	54	24.9	205	US-09-270-767-34638	Sequence 34638, A
10	54	24.9	205	US-09-270-767-49855	Sequence 49855, A
11	54	24.9	2539	US-09-413-814-42	Sequence 42, Appl
12	53	24.4	457	US-09-252-991A-32812	Sequence 32812, A
13	52	24.0	218	US-09-252-991A-18053	Sequence 18053, A
14	51	23.5	65	US-09-134-000C-5463	Sequence 5463, Appl
15	51	23.5	705	US-09-538-092-626	Sequence 626, Appl
16	51	23.5	1169	US-08-542-921-2	Sequence 2, Appl
17	51	23.5	1169	US-08-880-685-2	Sequence 2, Appl
18	51	23.5	1169	US-08-880-684-2	Sequence 2, Appl
19	50	23.0	228	US-09-248-796A-19282	Sequence 19282, A
20	50	23.0	353	US-09-270-767-39090	Sequence 39090, A
21	50	23.0	353	US-09-270-767-54307	Sequence 54307, A
22	49.5	22.8	208	US-09-107-532A-5399	Sequence 5399, Appl
23	49.5	22.8	475	US-09-543-681A-4698	Sequence 4698, Appl
24	49.5	22.8	523	US-09-252-991A-18693	Sequence 18693, A
25	49.5	22.8	1164	US-09-457-708-2	Sequence 2, Appl
26	49.5	22.8	1164	US-09-950-046A-2	Sequence 2, Appl
27	49.5	22.8	1164	US-09-976-594-989	Sequence 989, Appl

28	48.5	22.4	5032	4	US-09-538-092-979	Sequence 979, Appl
29	48	22.1	89	4	US-09-732-210-98	Sequence 98, Appl
30	48	22.1	96	4	US-09-513-999C-7299	Sequence 7299, Appl
31	48	22.1	1013	4	US-09-248-796A-18605	Sequence 18605, A
32	47.5	21.9	450	4	US-09-543-681A-5432	Sequence 5432, Appl
33	47	21.7	20	5	PCT-US95-07543-5	Sequence 5, Appl
34	47	21.7	74	4	US-09-134-000C-6669	Sequence 6669, Appl
35	47	21.7	91	4	US-09-732-210-195	Sequence 195, Appl
36	47	21.7	318	4	US-09-270-767-36225	Sequence 36225, A
37	47	21.7	318	4	US-09-270-767-51442	Sequence 51442, A
38	47	21.7	421	3	US-09-002-567B-1	Sequence 1, Appl
39	47	21.7	421	3	US-09-002-567B-3	Sequence 3, Appl
40	47	21.7	421	3	US-09-571-347-1	Sequence 1, Appl
41	47	21.7	421	3	US-09-571-347-3	Sequence 3, Appl
42	47	21.7	567	3	US-09-188-811-2	Sequence 2, Appl
43	47	21.7	600	4	US-09-270-767-42652	Sequence 42652, A
44	47	21.7	621	4	US-09-489-039A-10378	Sequence 10378, A
45	46.5	21.4	274	4	US-09-252-991A-29653	Sequence 29653, A

ALIGNMENTS

RESULT 1  
US-09-538-092-833  
; Sequence 833, Application US/09538092  
; Patent No. 6753314  
; GENERAL INFORMATION:  
; APPLICANT: Giot, Loic  
; APPLICANT: Mansfield, Traci A.  
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same  
; FILE REFERENCE: 15966-542  
; CURRENT APPLICATION NUMBER: US/09/538,092  
; CURRENT FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: 60/127,352  
; PRIOR FILING DATE: 1999-04-01  
; PRIOR APPLICATION NUMBER: 60/178,965  
; PRIOR FILING DATE: 2000-02-01  
; NUMBER OF SEQ ID NOS: 1387  
; SOFTWARE: CuraPatSeqformatter Version 0.9  
; SEQ ID NO 833  
; LENGTH: 354  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (0)..(0)  
; OTHER INFORMATION: Polypeptide Accession Number P00480  
US-09-538-092-833

Query Match 65.0%; Score 141; DB 4; Length 354;  
Best Local Similarity 69.0%; Pred. No. 4.2e-14;  
Matches 29; Conservative 4; Mismatches 9; Indels 0; Gaps 0;  
Qy 1 MLSNRLILNKALKAHKAHTSMVNRFRYKPVQSVOLKPRDL 42  
Db 1 MLFNRLILLNNAFRNGHFMVNRFRCGQLQNKVLKGRDL 42

RESULT 2  
US-08-373-190-29  
; Sequence 29, Application US/08373190  
; Patent No. 5851829  
; GENERAL INFORMATION:  
; APPLICANT: MARASCO, WAYNE  
; APPLICANT: HASELTINE, WILLIAM  
; TITLE OF INVENTION: METHOD OF INTRACELLULAR BINDING OF PROTEINS  
; NUMBER OF SEQUENCES: 79  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN  
; STREET: 130 WATER STREET  
; CITY: BOSTON  
; STATE: MA

```
/ COUNTRY: US
/ ZIP: 02109
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FASTSEQ Version 1.5
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/373,190
/ FILING DATE: 17-JAN-1995
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US93/06735
/ FILING DATE: 16-JUL-1993
/ ATTORNEY/AGENT INFORMATION:
/ NAME: RESNICK, DAVID S
/ REGISTRATION NUMBER: 34,235
/ REFERENCE/DOCKET NUMBER: 41956-PCT-US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617-523-3400
/ TELEFAX: 617-523-6440
/ TELEX: STRE UR 2002
/ INFORMATION FOR SEQ ID NO: 29:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 32 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: N-terminal
/ ORIGINAL SOURCE:
/ US-08-373-190-29

Query Match 42.4%; Score 92; DB 2; Length 32;
Best Local Similarity 61.3%; Pred. No. 1.4e-07;
Matches 19; Conservative 0; Indels 0; Gaps 0;

QY 1 MLSNRLILNKAALRKAKHTSMVNRFRYKPV 31
Db 1 MLFNLRXXLNNAAFRHGHFMVNRFCGQPL 31

RESULT 3
US-08-438-190A-29
/ Sequence 29, Application US/08438190A
/ Patent No. 5965371
/ GENERAL INFORMATION:
/ APPLICANT: MARASCO, WAYNE
/ APPLICANT: HASELTINE, WILLIAM
/ TITLE OF INVENTION: METHOD OF INTRACELLULAR BINDING OF
/ TITLE OF INVENTION: PROTEINS
/ NUMBER OF SEQUENCES: 78
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
/ ADDRESSEE: CUSHMAN
/ STREET: 130 WATER STREET
/ CITY: BOSTON
/ STATE: MASSACHUSETTS
/ COUNTRY: US
/ ZIP: 02109
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/438,190A
/ FILING DATE:
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: EISENSTEIN, RONALD I.

Query Match 42.4%; Score 92; DB 2; Length 32;
Best Local Similarity 61.3%; Pred. No. 1.4e-07;
Matches 19; Conservative 0; Indels 0; Gaps 0;

QY 1 MLSNRLILNKAALRKAKHTSMVNRFRYKPV 31
Db 1 MLFNLRXXLNNAAFRHGHFMVNRFCGQPL 31

RESULT 4
US-08-350-215-29
/ Sequence 29, Application US/08350215
/ Patent No. 6004940
/ GENERAL INFORMATION:
/ APPLICANT: MARASCO, WAYNE A.
/ APPLICANT: RICHARDSON, JENNIFER
/ TITLE OF INVENTION: INTRACELLULAR TARGETING OF ENDOGENOUS
/ TITLE OF INVENTION: PROTEINS
/ NUMBER OF SEQUENCES: 58
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
/ ADDRESSEE: CUSHMAN
/ STREET: 130 WATER STREET
/ CITY: BOSTON
/ STATE: MASSACHUSETTS
/ COUNTRY: US
/ ZIP: 02109
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/350,215
/ FILING DATE: 12-DEC-1994
/ CLASSIFICATION: 514
/ ATTORNEY/AGENT INFORMATION:
/ NAME: EISENSTEIN, RONALD I.
/ REGISTRATION NUMBER: 30628
/ REFERENCE/DOCKET NUMBER: 41956-CP3
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (617) 523-3400
/ TELEFAX: (617) 523-6440
/ TELEX: 200291 STRE UR
/ INFORMATION FOR SEQ ID NO: 29:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 32 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ US-08-350-215-29

Query Match 42.4%; Score 92; DB 3; Length 32;
Best Local Similarity 61.3%; Pred. No. 1.4e-07;
Matches 19; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLSNRLILNKAALRKAKHTSMVNRFRYKPV 31
Db 1 MLFNLRXXLNNAAFRHGHFMVNRFCGQPL 31
```

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/556,111
; FILING DATE: 21-Apr-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/438,190
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: EISENSTEIN, RONALD I.
; REGISTRATION NUMBER: 30628
; REFERENCE/DOCKET NUMBER: 41956
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 523-3400
; TELEFAX: (617) 523-6440
; TELEX: 200291 STRE UR
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 29:
;
US-09-556-111-29

Query Match 42.4%; Score 92; DB 3; Length 32;
Best Local Similarity 61.3%; Pred. No. 1.4e-07;
Matches 19; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLSNRLILNKAALRKAHTSMVRNFRYKGPV 31
Db 1 MLFNLRXXLNNAAFRGHFMVRNFRCGQPL 31

RESULT 7
US-09-270-767-56794
; Sequence 56794, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 56794
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-56794

Query Match 25.8%; Score 56; DB 4; Length 226;
Best Local Similarity 43.6%; Pred. No. 0.87;
Matches 17; Conservative 5; Mismatches 9; Indels 8; Gaps 2;

QY 2 LSNRLILNKAALR---KAHTSMVRNFRYKGPVQSQVL 37
Db 68 LHNRLILNHSQSRVKTKHTSVTYD-----PVGKRVLL 101

RESULT 8
US-09-270-767-41564
; Sequence 41564, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17

```

; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 41564  
; LENGTH: 602  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-09-270-767-41564

Query Match 25.8%; Score 56; DB 4; Length 602;  
Best Local Similarity 43.6%; Pred. No. 2.9;  
Matches 17; Conservative 5; Mismatches 9; Indels 8; Gaps 2;

Qy 2 LSNRIILNKAALR---KATTSWNRNFRYKPKVQSQVL 37  
| | | | | : | | | | | : | | | | | : | | | | |  
Db 444 LHNRIILNQLSRVKVKTHTSVTYD-----PVGKRVLL 477

RESULT 9  
US-09-270-767-34638  
; Sequence 34638, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 34638  
; LENGTH: 205  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
; FEATURE:  
; OTHER INFORMATION: Xaa means any amino acid  
US-09-270-767-34638

Query Match 24.9%; Score 54; DB 4; Length 205;  
Best Local Similarity 41.4%; Pred. No. 1.6;  
Matches 12; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

Qy 8 LLNKAALRKATTSWNRNFRYKPKVQSQVQ 36  
| | | | | : | | | | | : | | | | | : | | | | |  
Db 66 LINVAFPHKAYTILIRNLRFFSYKSHVK 94

RESULT 10  
US-09-270-767-49855  
; Sequence 49855, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 49855  
; LENGTH: 205  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
; FEATURE:  
; OTHER INFORMATION: Xaa means any amino acid  
US-09-270-767-49855

Query Match 24.9%; Score 54; DB 4; Length 205;  
Best Local Similarity 41.4%; Pred. No. 1.6;  
Matches 12; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

Qy 8 LLNKAALRKATTSWNRNFRYKPKVQSQVQ 36  
| | | | | : | | | | | : | | | | | : | | | | |  
Db 66 LINVAFPHKAYTILIRNLRFFSYKSHVK 94

## RESULT 11

US-09-413-814-42  
; Sequence 42, Application US/09413814  
; Patent No. 6225064  
; GENERAL INFORMATION:  
; APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH  
; APPLICANT: Bristol-Myers Squibb, Co.  
; APPLICANT: Beyer, Stefan  
; APPLICANT: Bloecker, Helmut  
; APPLICANT: Brandt, Petra  
; APPLICANT: Cino, Paul M  
; APPLICANT: Dougherty, Brian A  
; APPLICANT: Goldberg, Steven L  
; APPLICANT: Hofle, Gerhard  
; APPLICANT: Mueller, Joachim  
; APPLICANT: Reichenbach, Hans  
; TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or  
; TITLE OF INVENTION: heteropolyketide compounds  
; FILE REFERENCE: PCT/US 99/23535  
; CURRENT APPLICATION NUMBER: US/09/413,814  
; CURRENT FILING DATE: 1999-10-07  
; EARLIER APPLICATION NUMBER: DE 198 45 493.2  
; EARLIER FILING DATE: 1998-10-09  
; NUMBER OF SEQ ID NOS: 107  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 42  
; LENGTH: 2539  
; TYPE: PRT  
; ORGANISM: Sorangium cellulosum  
US-09-413-814-42

Query Match 24.9%; Score 54; DB 3; Length 2539;  
Best Local Similarity 35.5%; Pred. No. 37;  
Matches 11; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

Qy 12 AALRKAHTSMVRNFRYKPKVQSQVQLKPRDL 42  
| | | | | : | | | | | : | | | | | : | | | | |  
Db 1969 ASLSAHSSEPEPARHGRPALSSSEWVAPRNV 1999

RESULT 12  
US-09-252-991A-32812  
; Sequence 32812, Application US/09252991A  
; Patent No. 6551795

US-09-252-991A-32812  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 32812  
; LENGTH: 457  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (451)  
; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.  
US-09-252-991A-32812

Query Match 24.4%; Score 53; DB 4; Length 457;  
Best Local Similarity 34.4%; Pred. No. 6.3;  
Matches 11; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

Qy 7 ILLNKAALRKATSMVRNFRYKPKVQSQVQLK 38



```

; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 626
; LENGTH: 705
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: Polypeptide Accession Number YMR163C
US-09-538-092-626

Query Match 23.5%; Score 51; DB 4; Length 705;
Best Local Similarity 36.8%; Pred. No. 22;
Matches 14; Conservative 5; Mismatches 19; Indels 0; Gaps 0;

QY 5 LRILLNKAIRKAKHSMVNFYRGPVQSQVQLKPRDL 42
   || : |||| || : ||: ||: ||
DB 643 LRTAKSFELLRKAQASMSVKFGFKPLRDDAFLESRPL 680

RESULT 16
US-08-542-921-2
; Sequence 2, Application US/08542921
; Patent No. 5736514
; GENERAL INFORMATION:
; APPLICANT: IIZUKA, TOSHIHIKO
; APPLICANT: TAGAWA, MICHITO
; APPLICANT: ARAI, SATOSHI
; APPLICANT: NIIZEKI, MASATSUGU
; APPLICANT: MIYAKE, TOSHIRO
; TITLE OF INVENTION: NOVEL BACILLUS STRAIN AND HARMFUL
; TITLE OF INVENTION: ORGANISM CONTROLLING AGENTS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/542,921
; FILING DATE: 13-OCT-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 276082/94
; FILING DATE: 14-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 49-209-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1169 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

```







```

;
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-950-046A-2

Query Match      22.8%; Score 49.5; DB 4; Length 1164;
Best Local Similarity 29.8%; Pred. No. 73;
Matches 14; Conservative 9; Mismatches 19; Indels 5; Gaps 1;

QY 1 MLSNLRILLNKAALRKHAFTSMV-----RNFYRGKPVQSQVQLKPRDL 42
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 806 MIAELRIELKANKKVCHTELLLSQVSKLSNSESVOQQMEFLNRQL 852

RESULT 27
US-09-976-594-989
; Sequence 989, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 989
; LENGTH: 1164
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: incyte ID No. 6673549 4215034CD1
US-09-976-594-989

Query Match      22.8%; Score 49.5; DB 4; Length 1164;
Best Local Similarity 29.8%; Pred. No. 73;
Matches 14; Conservative 9; Mismatches 19; Indels 5; Gaps 1;

QY 1 MLSNLRILLNKAALRKHAFTSMV-----RNFYRGKPVQSQVQLKPRDL 42
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 806 MIAELRIELKANKKVCHTELLLSQVSKLSNSESVOQQMEFLNRQL 852

RESULT 28
US-09-538-092-979
; Sequence 979, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 979
; LENGTH: 5032
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number P21817
US-09-538-092-979

```

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/07543
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Perryman, David G.
REGISTRATION NUMBER: 33,438
REFERENCE/DOCKET NUMBER: 2200.0191
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404/688-0770
TELEFAX: 404/688-9880
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..20
OTHER INFORMATION: /note= "Description: mitochondrial
OTHER INFORMATION: localization signal"
PCT-US95-07543-5

Query Match 21.7%; Score 47; DB 5; Length 20;
Best Local Similarity 73.3%; Pred. No. 1.2;
Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MLSNLRILLNKAALR 15
   |||||: |||
Db 1 MLFNLRILLDDAAFR 15

RESULT 34
US-09-134-000C-6669
; Sequence 6669, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6669
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-6669

Query Match 21.7%; Score 47; DB 4; Length 74;
Best Local Similarity 39.4%; Pred. No. 6;
Matches 13; Conservative 5; Mismatches 7; Indels 8; Gaps 2

QY 5 LRILLNKAALRKHAHTSMVRNFR-----YGKPVQS 33
   |||: ||| |||
Db 13 LKILKQLLLSKAH----RNFTSPQVIGEPYGS 41

RESULT 35
US-09-732-210-195
; Sequence 195, Application US/09732210
; Patent No. 6573361
; GENERAL INFORMATION:
; APPLICANT: Bunkers, Greg J.
; APPLICANT: Liang, Jihong

```

APPLICANT: Mittanck, Cindy A.  
APPLICANT: Seale, Jeffrey W.  
APPLICANT: Wu, Yonnie S.  
TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use  
FILE REFERENCE: 38-21(15036)B  
CURRENT APPLICATION NUMBER: US/09/732,210  
CURRENT FILING DATE: 2000-12-07  
PRIOR APPLICATION NUMBER: US 60/169,513  
PRIOR FILING DATE: 1999-12-07  
PRIOR APPLICATION NUMBER: US 60/169,340  
PRIOR FILING DATE: 1999-12-07  
NUMBER OF SEQ ID NOS: 1753  
SEQ ID NO 195  
LENGTH: 91  
TYPE: PRT  
ORGANISM: Schizosaccharomyces pombe  
US-09-732-210-195

Query Match 21.7%; Score 47; DB 4; Length 91;  
Best Local Similarity 52.9%; Pred. No. 7.7;  
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 14 LRKAHTSMVRNFRYKGP 30  
DB 67 LKKVHRSFKNGFRSGKP 83

RESULT 36  
US-09-270-767-36225  
Sequence 36225, Application US/09270767  
Patent No. 6703491  
GENERAL INFORMATION:  
APPLICANT: Homburger et al.  
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
FILE REFERENCE: File Reference: 7326-094  
CURRENT APPLICATION NUMBER: US/09/270,767  
CURRENT FILING DATE: 1999-03-17  
NUMBER OF SEQ ID NOS: 62517  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 36225  
LENGTH: 318  
TYPE: PRT  
ORGANISM: Drosophila melanogaster  
FEATURE:  
OTHER INFORMATION: Xaa means any amino acid  
US-09-270-767-36225

Query Match 21.7%; Score 47; DB 4; Length 318;  
Best Local Similarity 29.4%; Pred. No. 36;  
Matches 10; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

QY 1 MLSNRLILNKAALRKHAHTSMVRNFRYKGPVQSQ 34  
DB 285 IISTLNKQNKAKQKKKTKLRSKCNKLIKNR 318

RESULT 37  
US-09-270-767-51442  
Sequence 51442, Application US/09270767  
Patent No. 6703491  
GENERAL INFORMATION:  
APPLICANT: Homburger et al.  
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
FILE REFERENCE: File Reference: 7326-094  
CURRENT APPLICATION NUMBER: US/09/270,767  
CURRENT FILING DATE: 1999-03-17  
NUMBER OF SEQ ID NOS: 62517  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 51442  
LENGTH: 318  
TYPE: PRT  
ORGANISM: Drosophila melanogaster  
FEATURE:

OTHER INFORMATION: Xaa means any amino acid  
US-09-270-767-51442

Query Match 21.7%; Score 47; DB 4; Length 318;  
Best Local Similarity 29.4%; Pred. No. 36;  
Matches 10; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

QY 1 MLSNRLILNKAALRKHAHTSMVRNFRYKGPVQSQ 34  
DB 285 IISTLNKQNKAKQKKKTKLRSKCNKLIKNR 318

RESULT 38  
US-09-002-567B-1  
Sequence 1, Application US/09002567B  
Patent No. 6001594  
GENERAL INFORMATION:  
APPLICANT: Lal, Preeti  
APPLICANT: Guegler, Karl J.  
APPLICANT: Corley, Neil C.  
TITLE OF INVENTION: HUMAN TESTIN  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Dr.  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/002,567B  
FILING DATE: December 31, 1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0454 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 421 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: NEUTFWT01  
CLONE: 338680  
US-09-002-567B-1

Query Match 21.7%; Score 47; DB 3; Length 421;  
Best Local Similarity 27.3%; Pred. No. 52;  
Matches 12; Conservative 10; Mismatches 16; Indels 6; Gaps 1;

QY 4 NLRILNKAALRKHAHTSMVRNFRYKGPVQSQ-----VOLKPRD 41  
DB 90 NWMLTNPVAAKNVSNITVTYEWAPPVQNALAQYMQMLPKE 133

RESULT 39  
US-09-002-567B-3  
Sequence 3, Application US/09002567B  
Patent No. 6001594  
GENERAL INFORMATION:  
APPLICANT: Lal, Preeti

APPLICANT: Guegler, Karl J.  
APPLICANT: Corley, Neil C.  
TITLE OF INVENTION: HUMAN TESTIN  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Dr.  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/002,567B  
FILING DATE: December 31, 1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0454 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 421 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 475210  
US-09-002-567B-3

Query Match 21.7%; Score 47; DB 3; Length 421;  
Best Local Similarity 27.3%; Pred. No. 52;  
Matches 12; Conservative 10; Mismatches 16; Indels 16; Gaps 1;  
QY 4 NLRILLNKAALRKAHTSMVRNFRYKGPVQSQ-----VOLKPRD 41  
DB 90 NVMILTNPVAAKNVSNITVTYEWAPPVQNALAQYQMPLPKE 133

RESULT 40  
US-09-571-347-1  
Sequence 1, Application US/09571347  
Patent No. 6358711  
GENERAL INFORMATION:  
APPLICANT: Lal, Preeti  
APPLICANT: Guegler, Karl J.  
APPLICANT: Corley, Neil C.  
TITLE OF INVENTION: HUMAN TESTIN  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Dr.  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/571,347

FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/002,567  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0454 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 421 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: NEUTIFWT01  
CLONE: 33860  
US-09-571-347-1  
Query Match 21.7%; Score 47; DB 3; Length 421;  
Best Local Similarity 27.3%; Pred. No. 52;  
Matches 12; Conservative 10; Mismatches 16; Indels 16; Gaps 1;  
QY 4 NLRILLNKAALRKAHTSMVRNFRYKGPVQSQ-----VOLKPRD 41  
DB 90 NVMILTNPVAAKNVSNITVTYEWAPPVQNALAQYQMPLPKE 133  
Search completed: December 18, 2004, 02:51:56  
Job time : 29.6667 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 18, 2004, 02:50:12 ; Search time 99.8214 Seconds  
(without alignments)  
154.144 Million cell updates/sec

Title: US-08-765-244-22

Perfect score: 217

Sequence: 1 MLSNRLILNKAAALRKAHTS.....NFRYKPVQSQVQLKPRDLC 43

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1589859 seqs, 357834939 residues

Total number of hits satisfying chosen parameters: 1589859

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	217	100.0	43	8	US-08-765-244-22
2	197	90.8	41	8	Sequence 22, Appl
3	92	42.4	32	13	US-08-765-244-1
4	92	42.4	32	13	Sequence 48, Appl
5	61	28.1	1286	10	US-10-052-942-54
6	61	28.1	1286	14	US-09-998-027-3
7	57	26.3	387	16	US-10-165-099-3
8	56	25.8	302	14	US-10-437-963-118548
9	53	24.4	268	16	US-10-369-493-3761
10	53	24.4	435	14	US-10-437-963-136742
11	53	24.4	593	15	US-10-369-493-22531
12	52.5	24.2	261	17	US-10-282-122A-78319
13	52.5	24.2	540	17	US-10-425-115-310467
					Sequence 7872, Ap

14	52	24.0	61	15	US-10-424-599-237920	Sequence 237920,
15	52	24.0	99	12	US-09-855-604-639	Sequence 639, App
16	51.5	23.7	468	14	US-10-369-493-4045	Sequence 4045, Ap
17	51	23.5	72	17	US-10-425-115-21280	Sequence 21280,
18	51	23.5	81	15	US-10-424-599-158049	Sequence 158049,
19	51	23.5	273	15	US-10-425-114-56810	Sequence 56810, A
20	51	23.5	296	17	US-10-425-115-209851	Sequence 209851,
21	51	23.5	598	15	US-10-282-122A-73330	Sequence 73330, A
22	50.5	23.3	383	14	US-10-369-493-17293	Sequence 17293, A
23	50.5	23.3	212	14	US-10-353-856-8	Sequence 8, Appl
24	50	23.0	414	14	US-10-156-761-7915	Sequence 7915, Ap
25	50	23.0	564	14	US-10-032-585-7102	Sequence 7102, Ap
26	49.5	22.8	118	16	US-10-767-701-38526	Sequence 38526, A
27	49.5	22.8	200	15	US-10-282-122A-57678	Sequence 57678, A
28	49.5	22.8	467	15	US-10-389-647-542	Sequence 542, App
29	49.5	22.8	807	16	US-10-437-963-174311	Sequence 174311,
30	49.5	22.8	1164	9	US-09-950-046A-2	Sequence 2, Appl
31	49	22.6	101	15	US-10-424-599-229228	Sequence 229228,
32	49	22.6	193	17	US-10-739-930-8796	Sequence 8796, Ap
33	49	22.6	261	16	US-10-437-963-145599	Sequence 145599,
34	49	22.6	296	16	US-10-437-963-129753	Sequence 129753,
35	49	22.6	355	14	US-10-369-493-21409	Sequence 21409, A
36	49	22.6	413	14	US-10-369-493-10261	Sequence 10261, A
37	49	22.6	446	15	US-10-282-122A-55126	Sequence 55126, A
38	49	22.6	479	14	US-10-369-493-8433	Sequence 8433, Ap
39	49	22.6	576	14	US-10-369-493-19977	Sequence 19977, A
40	49	22.6	672	16	US-10-437-963-119229	Sequence 119229,
41	49	22.6	1219	14	US-10-369-493-1885	Sequence 1885, Ap
42	48.5	22.4	71	15	US-10-424-599-214155	Sequence 214155,
43	48.5	22.4	75	9	US-09-864-761-46024	Sequence 46024, A
44	48.5	22.4	285	9	US-09-815-242-4869	Sequence 4869, Ap
45	48.5	22.4	295	15	US-10-282-122A-42529	Sequence 42529, A

#### ALIGNMENTS

RESULT 1

US-08-765-244-22  
; Sequence 22, Application US/08765244  
; Publication No. US2001000871A1  
; GENERAL INFORMATION:  
; APPLICANT: Seibel, Peter  
; APPLICANT: Seibel, Andrea  
; TITLE OF INVENTION: CHIMICAL PEPTIDE-NUCLEIC ACID  
; TITLE OF INVENTION: FRAGMENT, PROCESS FOR PRODUCING THE SAME AND ITS USES FOR  
; TITLE OF INVENTION: APPROPRIATELY INTRODUCING NUCLEIC ACIDS INTO CELL ORGANELLES  
; TITLE OF INVENTION: AND CELLS  
; FILE REFERENCE: 8484-0018-999  
; CURRENT APPLICATION NUMBER: US/08/765,244  
; CURRENT FILING DATE: 1997-10-30  
; PRIOR APPLICATION NUMBER: PCT/DE95/00775  
; PRIOR FILING DATE: 1995-06-11  
; PRIOR APPLICATION NUMBER: DE P 44 21 079.5  
; PRIOR FILING DATE: 1994-06-16  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 22  
; LENGTH: 43  
; TYPE: PRT  
; ORGANISM: Rattus rattus  
US-08-765-244-22

Query Match 100.0%; Score 217; DB 8; Length 43;  
Best Local Similarity 100.0%; Pred. No. 4.5e-24;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLSNRLILNKAAALRKAHTSMVRNFRYKPVQSQVQLKPRDLC 43

DB 1 MLSNRLILNKAAALRKAHTSMVRNFRYKPVQSQVQLKPRDLC 43

RESULT 2

```
US-08-765-244-1
; Sequence 1, Application US/08765244
; Publication No. US20010008771A1
; GENERAL INFORMATION:
; APPLICANT: Seibel, Peter
; APPLICANT: Seibel, Andrea
; TITLE OF INVENTION: CHEMICAL PEPTIDE-NUCLEIC ACID
; TITLE OF INVENTION: FRAGMENT, PROCESS FOR PRODUCING THE SAME AND ITS USES FOR
; TITLE OF INVENTION: APPROPRIATELY INTRODUCING NUCLEIC ACIDS INTO CELL ORGANELLES
; TITLE OF INVENTION: AND CELLS
; FILE REFERENCE: 8484-0018-999
; CURRENT APPLICATION NUMBER: US/08/765,244
; CURRENT FILING DATE: 1997-10-30
; PRIOR APPLICATION NUMBER: PCT/DE95/00775
; PRIOR FILING DATE: 1995-06-11
; PRIOR APPLICATION NUMBER: DE P 44 21 079.5
; PRIOR FILING DATE: 1994-06-16
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Rattus rattus
; FEATURE:
US-08-765-244-1

Query Match          90.8%; Score 197; DB 8; Length 41;
Best Local Similarity 95.3%; Pred. No. 3.5e-21;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 1;

QY 1 MLSNLRILLNKAALRKKAHTSMVRNFRYKGKPVQSQVQLKPRDLC 43
DB 1 MLSNLRILLNKAALRKKAHTSMVRNFRYKGKPVQS--QLKPRDLC 41

RESULT 3
US-10-061-395-48
; Sequence 48, Application US/10061395
; Publication No. US20020192675A1
; GENERAL INFORMATION:
; APPLICANT: Zauderer, Maurice
; APPLICANT: Smith, Ernest S.
; TITLE OF INVENTION: Methods of Identifying Regulator Molecules
; FILE REFERENCE: 1821.0080003
; CURRENT APPLICATION NUMBER: US/10/061,395
; CURRENT FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: 60/271,423
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/265,880
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/265,589
; PRIOR FILING DATE: 2001-02-02
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Heterologous signal sequence for the mitochondrial matrix
; NAME/KEY: MISC FEATURE
; LOCATION: (7)..(8)
; OTHER INFORMATION: May be any amino acid
; NAME/KEY: MISC FEATURE
; LOCATION: (32)..(32)
; OTHER INFORMATION: May be any amino acid
US-10-061-395-48

Query Match          42.4%; Score 92; DB 13; Length 32;
Best Local Similarity 61.3%; Pred. No. 5.6e-06;
Matches 19; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLSNLRILLNKAALRKKAHTSMVRNFRYKGPV 31
```

```
US-08-765-244-22.rapb
; Sequence 54, Application US/10052942
; Publication No. US20030104402A1
; GENERAL INFORMATION:
; APPLICANT: Zauderer, Maurice
; APPLICANT: Smith, Ernest
; APPLICANT: Wei, Chungwen
; TITLE OF INVENTION: Methods of Producing or Identifying Intrabodies in Eukaryotic Cells
; FILE REFERENCE: 1821.0090004
; CURRENT APPLICATION NUMBER: US/10/052,942
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: 60/298,095
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/271,422
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/263,200
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 60/263,225
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 154
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 54
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: signal sequence
; NAME/KEY: UNSURE
; LOCATION: (7)..(8)
; OTHER INFORMATION: Xaa may represent any amino acid
; NAME/KEY: UNSURE
; LOCATION: (32)..(32)
; OTHER INFORMATION: Xaa may represent any amino acid
US-10-052-942-54

Query Match          42.4%; Score 92; DB 14; Length 32;
Best Local Similarity 61.3%; Pred. No. 5.6e-06;
Matches 19; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLSNLRILLNKAALRKKAHTSMVRNFRYKGPV 31
DB 1 MLSNLRILLNKAALRKKAHTSMVRNFRYKGPV 31

RESULT 5
US-09-998-027-3
; Sequence 3, Application US/0998027
; Publication No. US20030093819A1
; GENERAL INFORMATION:
; APPLICANT: D'Andrea et al.
; TITLE OF INVENTION: Methods and Compositions for the
; TITLE OF INVENTION: Diagnosis and Treatment of Cancers Associated with Defective
; TITLE OF INVENTION: DNA Repair Mechanisms
; FILE REFERENCE: 2486/101
; CURRENT APPLICATION NUMBER: US/09/998,027
; CURRENT FILING DATE: 2001-11-02
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1286
; TYPE: PRT
; ORGANISM: A. thaliana
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(1286)
; OTHER INFORMATION: Plantfancd2
US-09-998-027-3
```

```

Query Match          28.1%; Score 61; DB 10; Length 1286;
Best Local Similarity 38.2%; Pred.No. 14;
Matches 13; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

QY      3  SNLRILINKAALRKAHTSMVRNFRYKPVQSOVQ 36
      |||| :| :||| :||| :||| :||| :||| :|||
Db      107  SNLRRLMLSSSTTKRDESLVRNLLLVSPILQDIO 140

RESULT 6
US-10-165-099-3
; Sequence 3, Application US/10165099
; Publication No. US20030188326A1
; GENERAL INFORMATION:
; APPLICANT: D'Andrea, Alan
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS OF CANCER SUSCEPTIBILITY
; FILE REFERENCE: 7032/2055
; CURRENT APPLICATION NUMBER: US/10/165,099
; CURRENT FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: US 09/998,027
; PRIOR FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: US 60/245,756
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 352
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1286
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-165-099-3

```

```

Query Match      28.1%; Score 61; DB 14; Length 1286;
Best Local Similarity 38.2%; Pred. No. 14;
Matches 13; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

Qy      3  SNRLFLNKAAALRKAHTSMVRNFRYGRPVQSQV 36
      ||||| : : : : : : : : : : : :
Db      107 SNLRMLSSSSSTTKRDESLVRLNLLVSPIDLIO 140

```

```

RESULT 7
US-10-437-963-118548
; Sequence 118548, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 118548
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(387)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_21849C.1.pep
US-10-437-963-118548

```

Query Match 26.3%; Score 57; DB 16; Length 387;

```

Best Local Similarity 35.9%; Pred. NO. 13;
Matches 14; Conservative 4; Mismatches 13; Indels 8; Gaps 1;

QY 2 LSNRLILNKAALRKAHTSMVRNFRYGRFPVQSQVQLKPR 40
   :||| :||| :||| :||| :||| :||| :||| :|||
Db 323 MSELR-----RDAHTQVYREQRWAKPTAEQLAADPR 353

RESULT 8
US-10-369-493-3761
; Sequence 3761, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 3761
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Neurospora crassa
US-10-369-493-3761

```

```

Query Match      25.8%; Score 56; DB 14; Length 302;
Best local Similarity 64.7%; Pred. No. 14;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY      12 AALRKAHTSMVRNFRYG 28
DB      271 AAMTKAHVSQMRWFRYG 287

```

```

RESULT 9
US-10-437-963-136742
; Sequence 136742, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 136742
; LENGTH: 268
; TYPE: PRI
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_38291C.1.pgp
US-10-437-963-136742

```

Query Match 24.4%; Score 53; DB 16; Length 268;  
Best Local Similarity 34.3%; Pred. No. 33;  
Matches 12; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

Qy 9 LNKAAALRKAMTSMVNRFRYKGPVQSQVLKPRDLG 43



US-10-739-930-7872

Query Match 24.2%; Score 52.5; DB 17; Length 540;  
Best Local Similarity 45.5%; Pred. No. 89;  
Matches 15; Conservative 4; Mismatches 11; Indels 3; Gaps 1;

QY 11 KAALRKA---HTSMVRNFRYKPKVQSQVOLKPR 40  
DB 506 EAALIRAIKPHAGALRNMYGKRVLSKACLKSR 538

RESULT 14

US-10-424-599-237920  
; Sequence 237920, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 237920  
; LENGTH: 61  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_56867C.1.pap  
US-10-424-599-237920

Query Match 24.0%; Score 52; DB 15; Length 61;  
Best Local Similarity 34.1%; Pred. No. 8.2;  
Matches 14; Conservative 5; Mismatches 4; Indels 18; Gaps 2;

QY 21 MVRNFRYK-----KP-----VOSQVOLKPRDLC 43  
DB 2 MYKNLRNGICCLISILWEDKPPPAERFVRVSQIKRPKILC 42

RESULT 15

US-09-855-604-639  
; Sequence 639, Application US/09855604  
; Publication No. US20040214165A1  
; GENERAL INFORMATION:  
; APPLICANT: GICQUEL, BRIGITTE  
; APPLICANT: FORTNOI, DENIS  
; APPLICANT: LIM, ENG-MONG  
; APPLICANT: PELICIC, VLADIMIR  
; APPLICANT: GUIGUENO, AGNES  
; APPLICANT: GOGUET DE LA SALMONIERE, YVES  
; TITLE OF INVENTION: POLYPEPTIDE NUCLEIC SEQUENCES EXPORTED FROM MYCOBACTERIA,  
; TITLE OF INVENTION: VECTORS COMPRISING SAME AND USES FOR DIAGNOSING AND  
; FILE REFERENCE: 03715.0062-01000  
; CURRENT APPLICATION NUMBER: US/09/855,604  
; CURRENT FILING DATE: 2001-05-16  
; PRIOR APPLICATION NUMBER: 09/485,536  
; PRIOR FILING DATE: 2000-02-14  
; PRIOR APPLICATION NUMBER: PCT/FR98/01813  
; PRIOR FILING DATE: 1998-08-14  
; PRIOR APPLICATION NUMBER: FR 97 10404  
; PRIOR FILING DATE: 1997-08-14  
; PRIOR APPLICATION NUMBER: FR 97 11325  
; PRIOR FILING DATE: 1997-09-11  
; NUMBER OF SEQ ID NOS: 935  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 639  
; LENGTH: 99  
; TYPE: PRT

; ORGANISM: Mycobacterium tuberculosis  
US-09-855-604-639

Query Match 24.0%; Score 52; DB 12; Length 99;  
Best Local Similarity 37.1%; Pred. No. 15;  
Matches 13; Conservative 8; Mismatches 12; Indels 2; Gaps 1;

QY 6 RILLNKAALRKHAHTSMVRNFRYKPKVQSQVOLKPR 40  
DB 53 RVILRCATRKANQSRATRLRPL--RIALRPR 85

RESULT 16

US-10-369-493-4045  
; Sequence 4045, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 4045  
; LENGTH: 468  
; TYPE: PRT  
; ORGANISM: Neurospora crassa  
US-10-369-493-4045

Query Match 23.7%; Score 51.5; DB 14; Length 468;  
Best Local Similarity 28.6%; Pred. No. 1.1e+02;  
Matches 12; Conservative 9; Mismatches 18; Indels 3; Gaps 1;

QY 2 LSNRLILNKAALRK---AHTSMVRNFRYKPKVQSQVOLKPR 40  
DB 269 IANSRVYQKSVAPKFEAFTEKMATIRGNPLDPTQMGQP 310

RESULT 17

US-10-425-115-212280  
; Sequence 212280, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 212280  
; LENGTH: 72  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_125201C.1.pap  
US-10-425-115-212280

Query Match 23.5%; Score 51; DB 17; Length 72;  
Best Local Similarity 39.3%; Pred. No. 14;  
Matches 11; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 9 LNKALRKHAHTSMVRNFRYKPKVQSQVQ 36

```
Db      2 LSKEALRFQTAIVHNTPLGEPAASHLQ 29
      |:| ||| |::| |::| |::| |::|
RESULT 18
US-10-424-599-158049
; Sequence 158049, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 158049
; LENGTH: 81
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_113738C.1.pep
US-10-424-599-158049
Query Match      23.5%; Score 51; DB 15; Length 81;
Best Local Similarity 47.8%; Pred. No. 16;
Matches 11; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY      13 ALRKAHTSMVRNFRYKPKVQSQV 35
      | | | | | | | | | | | | | | | |
Db      58 ASRNAKTGAIRQFRNGTPVRAGV 80
      | | | | | | | | | | | | | | | |

RESULT 19
US-10-425-114-56810
; Sequence 56810, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 56810
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLMO17039G01_FLI.pep
US-10-425-114-56810
Query Match      23.5%; Score 51; DB 15; Length 273;
Best Local Similarity 43.5%; Pred. No. 66;
Matches 10; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY      16 KAHTSMVRNFRYKPKVQSQV 38
      | | | | | | | | | | | | | | | |
Db      242 KPHTAVMQSNMYGKKVLSRICK 264
      | | | | | | | | | | | | | | | |

RESULT 20
US-10-425-115-209851
; Sequence 209851, Application US/10425115
; Publication No. US200400214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 209851
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_122978C.1.pep
US-10-425-115-209851
Query Match      23.5%; Score 51; DB 17; Length 296;
Best Local Similarity 43.5%; Pred. No. 73;
Matches 10; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY      16 KAHTSMVRNFRYKPKVQSQV 38
      | | | | | | | | | | | | | | | |
Db      265 KPHTAVMQSNMYGKKVLSRICK 287
      | | | | | | | | | | | | | | | |

RESULT 21
US-10-282-122A-73330
; Sequence 73330, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carx, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
```

```
; SEQ ID NO 73330
; LENGTH: 598
; TYPE: PRT
; ORGANISM: Salmonella paratyphi A
US-10-282-122A-73330

Query Match      23.5%; Score 51; DB 15; Length 598;
Best Local Similarity 33.3%; Pred. No. 1.7e+02;
Matches 14; Conservative 8; Mismatches 18; Indels 2; Gaps 1;

QY 1 MLSNRLILNKAALRKAKHTSMVRNF--RYGKPVQSOVOLKPR 40
Db 448 LITTFELQTEPLRLPALEASCOEFARFGFTVKLDYQLPPR 489

RESULT 22
US-10-369-493-17293
; Sequence 17293, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 17293
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Bacillus halodurans
US-10-369-493-17293

Query Match      23.3%; Score 50.5; DB 14; Length 383;
Best Local Similarity 48.3%; Pred. No. 1.2e+02;
Matches 14; Conservative 2; Mismatches 10; Indels 3; Gaps 1;

QY 2 LSNRLILNKAALRKAKHTSMVRNFYRYPK 30
Db 114 IQNKILLNSAATEKATGSPARG--GKP 139

RESULT 23
US-10-353-856-8
; Sequence 8, Application US/10353856
; Publication No. US20030215794A1
; GENERAL INFORMATION:
; APPLICANT: Kawakura, Yoshihiro
; APPLICANT: Jasenosky, Luke D.
; APPLICANT: Neumann, Gabriele
; APPLICANT: Wisconsin Alumni Research Foundation
; TITLE OF INVENTION: Filovirus Vectors and No. US20030215794A1 Infectious Filovirus-Bas
; FILE REFERENCE: 800.032US1
; CURRENT APPLICATION NUMBER: US/10/353,856
; CURRENT FILING DATE: 2003-01-29
; PRIOR APPLICATION NUMBER: US 60/353,972
; PRIOR FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 2212
; TYPE: PRT
; ORGANISM: Reston Ebola virus
US-10-353-856-8

Query Match      23.3%; Score 50.5; DB 14; Length 2212;
Best Local Similarity 25.6%; Pred. No. 9e+02;

Matches 11; Conservative 12; Mismatches 17; Indels 3; Gaps 1;

QY 1 MLSNRLILNKAALRKAKHTSMVRNFYRYPKPVQSOVOLKPRDL 43
Db 327 VINDRELISNRLKDYQOEKIRDFH---KILLQLQSPQFC 366

RESULT 24
US-10-156-761-7915
; Sequence 7915, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 7915
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-7915

Query Match      23.0%; Score 50; DB 14; Length 414;
Best Local Similarity 34.3%; Pred. No. 1.5e+02;
Matches 12; Conservative 5; Mismatches 18; Indels 0; Gaps 0;

QY 6 RILLNKAALRKAKHTSMVRNFYRYPKPVQSOVOLKPR 40
Db 294 RLQNTGAKPAAPARLTRAARYGGTGTVAEQPR 328

RESULT 25
US-10-032-585-7102
; Sequence 7102, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7102
; LENGTH: 564
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-032-585-7102

Query Match      23.0%; Score 50; DB 14; Length 564;
Best Local Similarity 29.1%; Pred. No. 2.2e+02;
Matches 16; Conservative 6; Mismatches 17; Indels 16; Gaps 1;

QY 2 LSNRLILNKAALR-----KAKHTSMVRNFYRYPKPVQSOVOLKPR 40
Db 424 LSDLLELLVIAAVRWTEKFELOTFNFNLAYTEQENVKFNFTGNAIVSSSTLSR 478
```

```

; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 57678
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Enterococcus faecium
US-10-282-122A-57678

Query Match      22.8%; Score 49.5; DB 15; Length 200;
Best Local Similarity 40.7%; Pred. No. 76;
Matches 11; Conservative 4; Mismatches 3; Indels 9; Gaps 1;

Qy    10 NKAALRKAAHTSMVNRFRYCKPVQSQQV 36
Db    51 NKAA-----QNFRYKGPFTPELE 68

RESULT 28
US-10-389-647-542
; Sequence 542, Application US/10389647
; Publication No. US20040033549A1
; GENERAL INFORMATION:
; APPLICANT: GREENBERG, E. Peter
; APPLICANT: SCHUSTER, Martin
; APPLICANT: LOSTROH, Candi
; TITLE OF INVENTION: QUORUM SENSING SIGNALING IN BACTERIA
; FILE REFERENCE: UIZ-038CP
; CURRENT APPLICATION NUMBER: US/10/389,647
; CURRENT FILING DATE: 2003-03-14
; PRIOR APPLICATION NUMBER: 09/653730
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/153022
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 710
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 542
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-389-647-542

Query Match      22.8%; Score 49.5; DB 15; Length 467;
Best Local Similarity 25.5%; Pred. No. 2.1e+02;
Matches 12; Conservative 11; Mismatches 9; Indels 15; Gaps 2;

Qy    6 RILLNKALRKAHTSMVRNFRY-----GKPVQSQQVOLK 38
Db    52 RMTANKSIEQAHTLLEQ-RYDLSDRPAKGASMTGRGPLQEGIRVK 97

RESULT 29
US-10-437-963-174311
; Sequence 174311, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 174311
; LENGTH: 807
; TYPE: PRT
; ORGANISM: Oryza sativa

```



```
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_72264C.1.pap
US-10-437-963-174311

Query Match      22.8%; Score 49.5; DB 16; Length 807;
Best Local Similarity 35.7%; Pred. No. 3.9e+02;
Matches 15; Conservative 9; Mismatches 15; Indels 3; Gaps 2;

QY 5 LRIL-LINKAALRAKHA--TSVNRNFRYKPKVQSQVQLKPRDL 43
Db 478 LRVLSLRASIHKVPDVSNLNLRLYLDAITRVKVIPLSLC 519

RESULT 30
US-09-950-046A-2
; Sequence 2, Application US/09950046A
; Patent No. US20020151701A1
; GENERAL INFORMATION:
; APPLICANT: Kwiatkowski, David J.
; Sampson, Julian R.
; Povey, Sue
; van Slegtenhorst, Marjon
; Halley, Dicky
; TITLE OF INVENTION: Compositions and Methods Based Upon the Tuberous
; Sclerosis-1 (TSC-1) Gene and Gene Product
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Vinson & Elkins
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.
; ZIP: 20004-1008
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/950,046A
; FILING DATE: 12-Sep-2001
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Sanzo, Michael A.
; REGISTRATION NUMBER: 36,912
; REFERENCE/DOCKET NUMBER: BR1331/42002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 639-6585
; TELEFAX: (202) 639-6604
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1164 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-950-046A-2

Query Match      22.8%; Score 49.5; DB 9; Length 1164;
Best Local Similarity 29.8%; Pred. No. 6e+02;
Matches 14; Conservative 9; Mismatches 19; Indels 5; Gaps 1;

QY 1 MLSNRLILNKALRKAHTSMV-----RNFRYKPKVQSQVQLKPRDL 42
Db 806 MIAELRIELKANNKVKCTELLLSQVSKLSNSVESVQQQMEFLNROL 852

RESULT 31
US-10-424-599-229228
; Sequence 229228, Application US/10424599
```

```
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 229228
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_49018C.1.pap
US-10-424-599-229228

Query Match      22.6%; Score 49; DB 15; Length 101;
Best Local Similarity 27.3%; Pred. No. 41;
Matches 9; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

QY 2 LSNRLILNKALRKAHTSMVNRNFRYKPKVQSQ 34
Db 1 LNELKLLQTKLSIMBKYSILHSGIFGPFDEQ 33

RESULT 32
US-10-739-930-8796
; Sequence 8796, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 8796
; LENGTH: 193
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: GLYMA-23APR03-C17728_1.p
US-10-739-930-8796

Query Match      22.6%; Score 49; DB 17; Length 193;
Best Local Similarity 32.0%; Pred. No. 87;
Matches 8; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

QY 10 NKAALRKAHTSMVNRNFRYKPKVQSQ 34
Db 32 SKIHMRNRHQKLLQHYRYPRETQSQ 56

RESULT 33
US-10-437-963-145599
; Sequence 145599, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
```

```

: FILE REFERENCE: 38-21(53221)B
: CURRENT APPLICATION NUMBER: US/10/437, 963
: CURRENT FILING DATE: 2003-05-14
: NUMBER OF SEQ ID NOS: 204966
: SEQ ID NO 145599
: LENGTH: 261
: TYPE: PRT
: ORGANISM: Oryza sativa
: FEATURE:
: OTHER INFORMATION: Clone ID: PAT_MRT4530.
US-10-437-963-145599

```

Query Match 22.6%; Score 49; DB 16; Length 261;  
Best Local Similarity 32.5%; Pred. No. 1.2e+02;  
Matches 13; Conservative 9; Mismatches 18; Indels

## RESULT 34

```

US-10-437-963-129753
; Sequence 129753, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid M
; TITLE OF INVENTION: Plants and Uses Th
; FILE REFERENCE: 38-21 (53221)B
; CURRENT APPLICATION NUMBER: US/10/437,9
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 129753
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(296)
; OTHER INFORMATION: unsure at all Xaa 1
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT45
US-10-437-963-129753

```

Query Match 22.6%; Score 49; DB 16; Length 296;  
Best Local Similarity 40.5%;  
Pred. No. 1.4e+02;  
Matches 15; Conservative 6; Mismatches 12; Indels

**Qy** 9 LNKAALRKAHTSMVRN---FRYCKPVQSQVQLKPRDL 42  
| : | : ||| : ||| : ||| : ||| :  
**Dd** 54 LSLALVNSSGGWIRSKRFVTGS-VTSVQLXPRL 89

RESULT 35

```

US-10-369-493-21409
/ Sequence 21409, Application US/10369493
/ Publication No. US2003023675A1
/ GENERAL INFORMATION:
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Hinkle, Gregory J.
/ APPLICANT: Slater, Steven C.
/ APPLICANT: Goldman, Barry S.
/ APPLICANT: Chen, Xianfeng
/ TITLE OF INVENTION: EXPRESSION OF MICROBIAL
/ TITLE OF INVENTION: PLANTS WITH IMPROVED

```

```

, FILE REFERENCE: 38-10(52052)B
, CURRENT APPLICATION NUMBER: US/10/369,493
, CURRENT FILING DATE: 2003-02-28
, PRIOR APPLICATION NUMBER: US 60/360,039
, PRIOR FILING DATE: 2002-02-21
, NUMBER OF SEQ ID NOS: 47374
, SEQ ID NO 21409
, LENGTH: 355
, TYPE: PRT
, ORGANISM: Archaeoglobus fulgidus
US-10-369-493-21409

```

Query Match	22.6%	Score 49;	DB 14;	Length 355;
Best Local Similarity	60.0%;	Pred. No. 1.8e+02;		
Matches	9;	Conservative	0;	Mismatches 6;
				Indels 0;
				Gaps 0;

Qy 15 RKAHTSMVRNFRYK 29  
||| | | | | |  
Db 312 RKLHRDFEVENFRYAK 326

RESULT 36

```

US-10-369-493-10261
; Sequence 10261, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 10261
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Cytophaga hutchinsonii
US-10-369-493-10261

```

Query Match	22.6%	Score 49;	DB 14;	Length 413;
Best Local Similarity	34.6%;	Pred. No. 2.1e+02;		
Matches	9;	Conservative	7;	Mismatches 10; Indels
				Gaps

Qy 3 SNLRILLNKAALRKAHTSMVRNFRYG 28  
: | : | | | | : : : | : |  
Db 378 ARKLTLNKAERKREVNVMCMNTFGFG 403

RESULT 37

```

US-10-282-122A-55126
; Sequence 55126, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of E
; FILE REFERENCE: ELI/TRA.034A

```

; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 55126  
; LENGTH: 446  
; TYPE: PRT  
; ORGANISM: Chlamydia trachomatis  
US-10-282-122A-55126

Query Match 22.6%; Score 49; DB 15; Length 446;  
Best Local Similarity 29.5%; Pred. No. 2.3e+02;  
Matches 13; Conservative 9; Mismatches 14; Indels 8; Gaps 1;  
QY 5 LRILLNKAALRKAAHTSMVNFYRG-----KPVQSQVQLKPR 40  
DB 226 LQPLVQKALRQVQAPFLAKSLKSGHKTVESYKPVETQALQFPQ 269

RESULT 38  
US-10-369-493-8433  
; Sequence 8433, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 8433  
; LENGTH: 479  
; TYPE: PRT  
; ORGANISM: Ralstonia metallidurans  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)-(479)  
; OTHER INFORMATION: unsure at all Xaa locations  
US-10-369-493-8433

Query Match 22.6%; Score 49; DB 14; Length 479;  
Best Local Similarity 45.0%; Pred. No. 2.5e+02;  
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;  
QY 18 HTSMVNFYRGKPVQSQVQL 37

DB 309 HRSIRDNLRYGAPDSTAEEL 328  
RESULT 39  
US-10-369-493-19977  
; Sequence 19977, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 19977  
; LENGTH: 576  
; TYPE: PRT  
; ORGANISM: NO. US20030233675Altoc punctiforme  
US-10-369-493-19977

Query Match 22.6%; Score 49; DB 14; Length 576;  
Best Local Similarity 40.0%; Pred. No. 3.1e+02;  
Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;  
QY 19 TSMVNFYRGKPVQSQVQLK 38  
DB 423 TTIKNIRYGDPISEGEQIE 442

RESULT 40  
US-10-437-963-119229  
; Sequence 119229, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 119229  
; LENGTH: 672  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_22466C.1.pep  
US-10-437-963-119229

Query Match 22.6%; Score 49; DB 16; Length 672;  
Best Local Similarity 25.5%; Pred. No. 3.7e+02;  
Matches 13; Conservative 9; Mismatches 19; Indels 10; Gaps 1;  
QY 1 MLSNRLILNKAALRKAAHTSMVR-----NFRYKPKVQSQVQLKPRD 41  
DB 375 ILKRLHIYKAKACLEIHSILRKHLRLEHFDPSIQDKSEDEIDAKKD 425

Search completed: December 18, 2004, 03:07:06  
Job time : 100.821 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 18, 2004, 02:37:05 ; Search time 24.0595 Seconds  
(without alignments)  
171.962 Million cell updates/sec

Title: US-08-765-244-22  
Perfect score: 217  
Sequence: 1 MLSNRIILNKAALRAHTS.....NFRYKPVQSVQLKPRDLIC 43

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	199	91.7	354	1 OWRT	ornithine carbamoy
2	164	75.6	354	1 OWMS	ornithine carbamoy
3	141	65.0	354	1 OWHU	ornithine carbamoy
4	119	54.8	41	2 I52779	ornithine transcar
5	61	28.1	1286	2 B71413	hypothetical prote
6	60	27.6	351	2 F90409	GTP binding conser
7	57	26.3	354	2 JE0309	ornithine carbamoy
8	55.5	25.6	454	2 A86345	F16P4.13 protein -
9	55	25.3	350	2 A48421	ornithine transcar
10	54	24.9	300	2 T32702	hypothetical prote
11	54	24.9	457	2 AD0950	oxygen-independent
12	54	24.9	1225	2 A56514	chromokinesin - ch
13	53.5	24.7	425	2 I40646	sensor-like protei
14	53	24.4	435	2 S31290	cyclin B5 - Yeast
15	53	24.4	593	2 AB0239	nitrate/nitrite se
16	53	24.4	593	2 T47000	nitrate/nitrite se
17	52.5	24.2	185	2 B81708	translation elonga
18	52	24.0	213	2 B83182	probable two-compo
19	52	24.0	335	2 T33457	hypothetical prote
20	52	24.0	598	1 RGEQNX	nitrate/nitrite se
21	52	24.0	598	2 P85702	nitrate/nitrite se
22	52	24.0	598	2 G90844	nitrate/nitrite se
23	51.5	23.7	260	2 T02333	hypothetical prote
24	51	23.5	506	2 T50211	WD-repeat protein
25	51	23.5	598	2 AP0648	nitrate/nitrite se
26	51	23.5	633	2 S76749	hypothetical prote
27	51	23.5	705	2 S54521	probable membrane
28	51	23.5	1779	2 T23130	hypothetical prote
29	50.5	23.3	138	2 P81900	hypothetical prote

30	50.5	23.3	383	2	B83922	short-chain-specif
31	50.5	23.3	429	2	E90267	hypothetical prote
32	50	23.0	178	2	H86026	hypothetical prote
33	50	23.0	178	2	F91180	hypothetical prote
34	50	23.0	202	2	F40590	motB homolog lafu
35	50	23.0	338	2	H75127	CAAX prenyl protei
36	50	23.0	386	2	F72773	probable molybdopt
37	50	23.0	616	2	A58947	signal recognition
38	50	23.0	662	2	AB1979	calcium-dependent
39	49.5	22.8	467	2	G83266	cytochrome c FA303
40	49.5	22.8	864	2	B90395	purine NTPase lim
41	49.5	22.8	1164	2	T03814	tumor suppressor p
42	49	22.6	176	2	S57240	18c protein (clone
43	49	22.6	301	2	T33068	hypothetical prote
44	49	22.6	355	1	B69518	GTP-binding protei
45	49	22.6	446	2	C81719	conserved hypothe

ALIGNMENTS

RESULT 1

OWRT  
ornithine carbamoyltransferase (EC 2.1.1.3.3) precursor - rat  
N:Alternate names: citrulline phosphorylase; ornithine transcarbamylase  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 28-Feb-1986 #sequence revision 28-Feb-1986 #text change 09-Jul-2004  
C:Accession: A00563; A28042; A23090; S02466; I52976; I67609; I53457  
R:Takiguchi, M.; Miura, S.; Mori, M.; Tatibana, M.; Nagata, S.; Kaziro, Y.  
Proc. Natl. Acad. Sci. U.S.A. 81, 7412-7416, 1984  
A>Title: Molecular cloning and nucleotide sequence of cDNA for rat ornithine carbamoyltransferase  
A:Reference number: A00563; MUID:85063800; PMID:6095294  
A:Accession: A00563  
A:Molecule type: mRNA  
A:Residues: 1-354 <TAKI>  
A:Cross-references: UNIPROT:P00481; GB:K03040; NID:G205873; PIDN:AAA41768.1; PID:G205874  
R:Takiguchi, M.; Murakami, T.; Miura, S.; Mori, M.  
Proc. Natl. Acad. Sci. U.S.A. 84, 6136-6140, 1987  
A>Title: Structure of the rat ornithine carbamoyltransferase gene, a large, X chromosome-  
A:Reference number: A28042; MUID:87317609; PMID:3476935  
A:Accession: A28042  
A:Molecule type: DNA  
A:Residues: 1-354 <TAKE>  
A:Cross-references: GB:M16933; GB:J02957; NID:G205884; PIDN:AAA41769.1; PID:G205886  
R:Kraus, J.P.; Hodges, P.E.; Williamson, C.L.; Horwich, A.L.; Kalousek, F.; Williams, K.F.  
Nucleic Acids Res. 13, 943-952, 1985  
A>Title: A cDNA clone for the precursor of rat mitochondrial ornithine transcarbamylase:  
A:Reference number: A23090; MUID:85215524; PMID:3839075  
A:Accession: A23090  
A:Molecule type: mRNA  
A:Residues: 1-38, 'P', 40-240, 'S', 242-354 <KRA>  
A:Cross-references: GB:X01976  
R:Aoki, Y.; Sunaga, H.; Suzuki, K.T.  
Biochem. J. 250, 735-742, 1988  
A>Title: A cadmium-binding protein in rat liver identified as ornithine carbamoyltransferase  
A:Reference number: S02466; MUID:88268748; PMID:3390141  
A:Accession: S02466  
A:Molecule type: protein  
A:Residues: 33-56;293-302;307-317;322-329 <AOK>  
R:McIntyre, P.; Graf, L.; Mercer, J.F.B.; Wake, S.A.; Hudson, P.J.; Hoogenraad, N.  
DNA 4, 147-156, 1985  
A>Title: The primary structure of the imported mitochondrial protein, ornithine transcarbamoyltransferase  
A:Reference number: I52976; MUID:85203360; PMID:3838931  
A:Accession: I52976  
A>Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-354 <RES>  
A:Cross-references: GB:M11266; NID:G205871; PIDN:AAA41767.1; PID:G205872  
R:McIntyre, P.; Graf, L.; Mercer, J.; Peterson, G.; Hudson, P.J.; Hoogenraad, N.  
FEBS Lett. 177, 41-46, 1984  
A>Title: A highly basic N-terminal extension of the mitochondrial matrix enzyme ornithin  
A:Reference number: I53457; MUID:85051832; PMID:6548714  
A:Accession: I67609

Matches 35; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 MLSNLRILLNAALRKAKHTSMVRNFRYPKPVQSQVLKPRDL 42  
||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 1 MLSNLRILLNNAAALKGHTSVRRHPWCGKPVSQVQLKGRDL 42

RESULT 3

OWHU

N;ornithine carbamoyltransferase (EC 2.1.1.3) precursor - human

N;Alternate names: citrulline phosphorylase; ornithine transcarbamylase

C;Species: Homo sapiens (man)

C;Date: 28-Feb-1986 #sequence\_revision 31-Mar-1993 #text\_change 09-Jul-2004

C;Accession: A41444; B41444; A00562; I38078; JC4672; I59039; I54377

R;Hata, A.; Tsuzuki, T.; Shimada, K.; Takiguchi, M.; Mori, M.; Matsuda, I.  
J. Biochem. 103, 302-308, 1988

A;Title: Structure of the human ornithine transcarbamylase gene.

A;Reference number: A41444; MUID:88227905; PMID:2836378

A;Accession: A41444

A;Molecule type: DNA

A;Residues: 1-354 <HAT>

A;Cross-references: UNIPROT:P00480; GB:D00230; NID:g219957; PIDN:BAA00161.1; PID:g219959

A;Accession: B41444

A;Molecule type: mRNA

A;Residues: 1-354 <HA2>

A;Cross-references: GB:D00230; NID:g219957; PIDN:BAA00161.1; PID:g219959

R;Horwich, A.L.; Fenton, W.A.; Williams, K.R.; Kalousek, F.; Kraus, J.P.; Doolittle, R.F.  
Science 224, 1068-1074, 1984

A;Title: Structure and expression of a complementary DNA for the nuclear coded precursor

A;Reference number: A00562; MUID:84196410; PMID:6372096

A;Accession: A00562

A;Molecule type: mRNA

A;Residues: 1-100,'P','102-110','P','112-192','CF','195-269','R','271-354 <HOR>

A;Cross-references: GB:D00230

R;Hata, A.; Tsuzuki, T.; Shimada, K.; Takiguchi, M.; Mori, M.; Matsuda, I.  
J. Biochem. 100, 717-725, 1986

A;Title: Isolation and characterization of the human ornithine transcarbamylase gene: st

A;Reference number: I38078; MUID:87057134; PMID:3782067

A;Accession: I38078

A>Status: translation not shown

A;Molecule type: DNA

A;Residues: 1-26 <RES>

A;Cross-references: EMBL:X04443; NID:g35162; PIDN:CAA28039.1; PID:g35163

R;Wheelier, V.C.; Prodmou, C.P.; Pearl, L.H.; Williamson, R.; Coutelle, C.  
Gene 169, 251-255, 1996

A;Title: Synthesis of a modified gene encoding human ornithine transcarbamylase for exper

etic defect.

A;Reference number: JC4672; MUID:96194812; PMID:8647457

A;Accession: JC4672

A;Molecule type: DNA

A;Residues: 'M','33-100','P','102-110','P','112-192','CF','195-269','R','271-354 <WHE>

A;Note: this report represents a synthetic gene designed for expression in (rather than t

R;Horwich, A.L.; Kalousek, F.; Rosenberg, L.E.  
Proc. Natl. Acad. Sci. U.S.A. 82, 4930-4933, 1985

A;Title: Arginine in the leader peptide is required for both import and proteolytic cleav

A;Reference number: I59039; MUID:85270440; PMID:3895227

A;Accession: I59039

A>Status: translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-36 <RES>

A;Cross-references: GB:M11235; NID:g189408; PIDN:AAA59976.1; PID:g189409

R;Gilbert-Dussardier, B.; Rabier, D.; Strautnieks, S.; Segues, B.; Bonnefont, J.P.; Munnu

Hum. Mol. Genet. 3, 831-832, 1994

A;Title: A novel arginine (245) to glutamine change in exon 8 of the ornithine carbamoyl

c pattern.

A;Reference number: I54377; MUID:94362689; PMID:8081373

A;Accession: I54377

A>Status: translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 269-276,'Q','278-289 <RE3>

A;Cross-references: GB:S73640; NID:g688001; PIDN:AAB31859.1; PID:g688002

A;Note: This sequence represents a disease defect in ornithine carbamoyltransferase

C;Comment: The active enzyme is a dimer of identical chains with one tightly bound zinc

e synthesis of UMP.  
C:Comment: The active enzyme catalyzes the condensation of carbamoyl phosphate and ornithine to form citrulline.  
C:Genetics:

**A:Gene: GDB:OTC**

A; Cross-references: GDB:119468; OMIM:311250  
A; Sup position: Xp21.1-Xp21.1  
C; Superfamily: ornithine carbamoyltransferase;  
C; Keywords: arginine biosynthesis; homotrimer;  
F; 1-2/Domain: transit peptide (mitochondrion);  
F; 3-354/Product: ornithine carbamoyltransferase #status predicted <RNP>  
F; 40-342/Domain: aspartate/ornithine carbamoyltransferase homolog <ACT>

Query Match	65.0%	Score 141;	DB 1;	Length 354;
Best Local Similarity	69.0%;	Pred. No. 9.9e-12;		
Matches 29;	Conservative	4;	Mismatches 9;	Indels 0;
				Gaps 0;

QY 1 MLNLRLILNKAALRKAHTSMVRNFRYGKPVQSQVLKPRDL 42

Db 1 MLFNLRLILNNAAFRNCHNFMVRNFRCGOPIONKQVLKGRDL 42

```

RESULT 4
152779
ornithine transcarbamylase peptide - rat (fragment)
C:Species: Rattus sp. (rat)
C:Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 18-Jun-1999
C:Accession: 152779
R:Horwich, A L.; Kalousek, P.; Fenton, W A.; Pollock, P A.; Rosenberga, L. E.

```

Cell 44, 451-459, 1986  
 A:Rothblat, A.G.; Kadosuek, F.; Feilich, W.A.; Follock, K.A.; Rosenberg, L.E.  
 A>Title: Targeting of pre-ornithine transcarbamylase to mitochondria: Definition of crith  
 A:Reference number: 152779; PMID:86106223; PMID:3943133  
 A:Accession: 152779  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-41 <RES>  
 A:Cross-references: GB:M12583; NID:g205887; PIDN:AAA41770.1; PID:g205888  
 C:Superfamily: ornithine carbamoyltransferase; aspartate/ornithine carbamoyltransferase

Query Match	54.8%	Score 119;	DB 2;	Length 41;
Best Local Similarity	66.7%	Pred. NO. 1.1e-09;		
Matches 24;	Conservative	4;	Mismatches 8;	Indels 0;
Gaps 0;				

**Qy**

1 MSLNLRIILNKAALRKAHTSMVRNFYRKPVQSQQV 36  
|| || || || || || | || || || || : : :

**Dd**

1 MLFNLRLINAAFGHFNFMVRNFCGLOPKVO 36  
| | | | | | | | | | | | | | | | | |

RESULT 5

B71413

hypothetical protein dl3525w - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

A:Variety: Columbia

C:Date: 03-Aug-1998 #sequence\_revision 03-Aug-1998 #text\_change 09-Jul-2004

C:Accession: B71413

R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirks, P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenegger, T.; Pohl, T.M.; Terryn, N.; Gielanavagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B. Nature 391, 485-488, 1998

A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech, A.; Moeres, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Auerhoft, A.; Chalmatzis, N. C.; Chalmatzis, N. C.

A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thaliana

A:Reference number: A71400; MUID:98121113; PMID:9461215

A:Accession: B71413

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-1286 <BEV>

A:Cross-references: UNIPROT:O23351; GS:Z97337; NID:g2244829; PID:g2244854

C:Genetics:

A;Map position: 4COP9-4G3845  
C;Superfamily: Arabidopsis thaliana hypothetical protein dl3525w

Query Match	28.1%	Score 61:	DB 2:	Length 1286:
-------------	-------	-----------	-------	--------------

Query Match	28.1%	Score 61:	DB 2:	Length 1286:
-------------	-------	-----------	-------	--------------

Best Local Similarity 38.2%; Pred. No. 5.1;  
Matches 13; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

Qy 3 SNLPIILLKAAALRKAHTSMVRNFRYGPVQSQVQ 36  
||||| : : : | : ||| : | : |  
Db 107 SNLRMLSSSTTKRDESLVRNLLLVSPQLDIO 140

RESULT 6  
F90409  
GTP binding conserved hypothetical protein SSO2385 [imported] - Sulfolobus solfataricus  
C:Species: Sulfolobus solfataricus  
C>Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 09-Jul-2004  
C:Accession: F90409  
R:She, Q.; Singh, R. K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M. J.; Chan  
Jong, I.; Jeffries, A. C.; Kozera, C. J.; Medina, N.; Peng, X.; Thi-Ngoc, H. P.; Redder, R.  
arrett, R. A.; Ragan, M. A.; Sensen, C. W.; Van der Oost, J.  
submitted to GenBank, April 2001  
A:Description: Sulfolobus solfataricus complete genome.

A;Description: *Sulfolobus solfataricus* Complete genome.  
A;Reference number: A99139  
A;Accession: F90409  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-351 <KUR>  
A;A,cross-references: UNIPROT:Q97W5; GB:AE006641; NID:g13815697; PIDN:AAK42533.1; GSPDB:  
C;Genetics:  
A;Gene: SSO2385

```

Query Match      27.6%; Score 60; DB 2; Length 351;
Best Local Similarity 34.6%; Pred. No. 1.7;
Matches 18; Conservative 8; Mismatches 10; Indels 16; Gaps 14
QY 7 ILLNKAA-----LRKAHTSMVRNFRY-----GKPVQSQVQ-----LKPRDL 42
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Dh 294 LTLKKGSFVLIVAKTILHSSLAERFYRVAUVWGKVXFGOGKVPSPSHLREDPI 345

```

RESULT 7  
JE0309  
ornithine carbamoyltransferase (EC 2.1.3.3) - chicken  
C/Species: Gallus gallus (Chicken)  
C/Date: 05-feb-1999 #sequence\_revision 05-Feb-1999 #text\_change 09-Jul-2004  
C/Accession: JE0309  
R/Shinogari, T.; Kono, M.; Mannen, H.; Mizutani, M.; Tsuji, S.  
J. Biochem. 124, 962-971, 1998  
A/Title: Chicken ornithine transcarbamoylase gene, structure, and  
A/Reference number: JE0309; MUID:99011321; PMID:9792920

A:Accession: JF0309  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-354 <SH1>  
A:Cross-references: UNIPROT:Q9VHY9; GB:AF065629; NID:G4218928; PIDM:AA01  
C:Superfamily: ornithine carbamoyltransferase; aspartate/ornithine carbamoyltransferase  
C:Keyword: transferase  
F:/40-342/Domain: aspartate/ornithine carbamoyltransferase homology <AOC>

Query Match 26.3%; Score 57; DB 2; Length 354;  
Best Local Similarity 40.5%;  
Pred. No. 4.5;  
Matches 17; Conservative 4; Mismatches 21; Indels 0; Gaps 0;

[illegible]

RESULT 8  
F16P4.13 protein - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001  
C/Accession: A86345  
R/Theologig, A.; Ecker, J.R.; Palm, C.J.; Federspiel,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.;  
Kane, S.M.; Karp, P.D.; Kornberg, J.R.; Kravchenko,  
L.; Lee, D.C.; Lin, Y.-D.; Liu, X.; Luo, R.; Ma,  
M.; McCombs, G.; Miller, R.; Morris, T.; Nelson,  
S.; O'Brien, A.D.; O'Neil, J.P.; Park, S.H.; Peck,  
J.L.; Peterson, J.L.; Quinn, W.A.; Schaefer,  
A.E.; Scheraga, H.A.; Sheng, Q.; Smith,  
T.F.; Sprague, A.M.; Stange, C.; Struhl,  
K.; Tamkun, J.W.; Tenenbaum,  
J.B.; Treisman, R.; Venter, A.; Wang,  
W.; Weisburger, S.; White,











C:Superfamily: Saccharomyces cerevisiae probable membrane protein YMR163C

C:Keywords: transmembrane protein

F:215-231/Domain: transmembrane #status predicted <TM1>

F:415-431/Domain: transmembrane #status predicted <TM2>

Query Match 23.5%; Score 51; DB 2; Length 705;

Best Local Similarity 36.8%; Pred. No. 65;

Matches 14; Conservative 5; Mismatches 19; Indels 0; Gaps 0;

QY 5 LRILLKAAALRKAAHTSMVRNFRYKGPQVQSQVLKPRDL 42

DB 643 LRTAKSFELLRKAAQASMSVKFGFQPLRDDAFLESRPL 680

RESULT 28

T23130

hypothetical protein T28B8.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C:Accession: T23130; T25403

R:White, S.

submitted to the EMBL Data Library, December 1997

A:Reference number: Z19690

A:Accession: T23130

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1779 <WIL>

A:Cross-references: UNIPROT:O18150; EMBL:AL021066; PIDN:CAA15925.1; GSPDB:GN00019; CESP:

A:Experimental source: clone H31B20

R:White, S.

submitted to the EMBL Data Library, October 1996

A:Reference number: Z20029

A:Accession: T25403

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1779 <W12>

A:Cross-references: EMBL:Z81133; PIDN:CAB03445.1; GSPDB:GN00019; CESP:T28B8.4

A:Experimental source: clone T28B8

C:Genetics:

A:Gene: CESP:T28B8.4

A:Map position: 1

A:Introns: 161/2; 223/2; 309/3; 332/2; 547/3; 603/1; 657/3; 745/2; 802/1; 856/1; 1031/1;

C:Superfamily: Caenorhabditis elegans hypothetical protein T28B8.3

Query Match 23.5%; Score 51; DB 2; Length 1779;

Best Local Similarity 46.7%; Pred. No. 1.8e+02;

Matches 14; Conservative 4; Mismatches 8; Indels 4; Gaps 1;

QY 1 MSLNRLIL----LNKAALRKAAHTSMVRNFR 26

DB 1718 ILKPLRLILSSSTSSNKAQVKKQTWTETVREPR 1747

RESULT 29

F81900

hypothetical protein NMA1316 [imported] - Neisseria meningitidis (strain Z2491 serogroup

C:Species: Neisseria meningitidis

C:Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 09-Jul-2004

C:Accession: F81900

R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel

R:Parkhill, S.; Jagsels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,

Nature 404, 502-506, 2000

A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.

A:Reference number: A81775; MUID:20222556; PMID:10761919

A:Accession: F81900

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-138 <PAR>

A:Cross-references: UNIPROT:Q9JUG9; GB:AL162755; GB:AL157959; NID:g7379742; PIDN:CAB8456

A:Experimental source: serogroup A, strain Z2491

C:Genetics:

A:Gene: NMA1316

Query Match 23.3%; Score 50.5; DB 2; Length 138;  
Best Local Similarity 30.4%; Pred. No. 13;  
Matches 14; Conservative 5; Mismatches 14; Indels 13; Gaps 1;

QY 6 RILLNKAALRKAAHTSMVNRFRY-----GKPVQSQVQLK 38  
DB 26 RYLLMRRLSETMHTAVKLNFRYAGRPKWGLKYRDGKPLSDSGRLK 71

RESULT 30  
B83922  
short-chain-specific acyl-CoA dehydrogenase BH2178 [imported] - Bacillus halodurans (strain  
C:Species: Bacillus halodurans  
C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 09-Jul-2004  
C:Accession: B83922  
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira-  
Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
A:Reference number: A83650; MUID:20512582; PMID:11058132  
A:Accession: B83922  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-383 <STO>  
A:Cross-references: UNIPROT:Q9KAV8; GB:AP001514; GB:BA000004; NID:G10174613; PIDN:BA0058  
A:Experimental source: strain C-125  
C:Genetics:  
A:Gene: BH2178

Query Match 23.3%; Score 50.5; DB 2; Length 383;  
Best Local Similarity 48.3%; Pred. No. 39;  
Matches 14; Conservative 2; Mismatches 10; Indels 3; Gaps 1;

QY 2 LSNRLIILNKAALRKAAHTSMVNRFRYKGP 30  
DB 114 IQNKQILLNSATEKATGSPARG---GKP 139

RESULT 31  
E90267  
hypothetical protein trpB-like [imported] - Sulfolobus solfataricus  
C:Species: Sulfolobus solfataricus  
C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 09-Jul-2004  
C:Accession: E90267  
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-  
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, R.;  
arrest, R.A.; Ragan, M.A.; Sensesen, C.W.; Van der Oost, J.  
submitted to GenBank, April 2001  
A:Description: Sulfolobus solfataricus complete genome.  
A:Reference number: A99139  
A:Accession: E90267  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-429 <KUR>  
A:Cross-references: UNIPROT:Q97TX6; GB:AE006641; NID:G13814334; PIDN:AAK41396.1; GSPDB:G  
C:Genetics:  
A:Gene: trpB-like  
C:Superfamily: tryptophan synthase beta chain; tryptophan synthase beta chain homology

Query Match 23.3%; Score 50.5; DB 2; Length 429;  
Best Local Similarity 35.6%; Pred. No. 44;  
Matches 16; Conservative 4; Mismatches 16; Indels 9; Gaps 2;

QY 7 ILLNKAALR-KAHTSMVNRFRYKGPVQ-----SQVQLKPRDL 42  
DB 143 VALASALFRMKAHIFWVNTSYAKPYRKYMMQMYGAEVHPSPSDL 187

RESULT 32  
H86026  
hypothetical protein Z4949 [imported] - Escherichia coli (strain O157:H7, substrain EDL5  
C:Species: Escherichia coli  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004  
C:Accession: H86026

R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,  
Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dialanta, E.; Potamouis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: A85480; MUID:21074935; PMID:11206551  
A:Accession: H86026  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-178 <STO>  
A:CROSS-references: UNIPROT:Q8X5L6; GB:AE005174; NID:g12518253; PIDN:AAGS8676.1; GSPDB:G  
C:Experimental source: strain O157:H7, substrain EDL933  
C:Genetics:  
A:Gene: Z4949

Query Match                23.0%; Score 50; DB 2; Length 178;  
Best Local Similarity      33.3%; Pred. No. 20;  
Matches     11; Conservative          6; Mismatches     16; Indels     0; Gaps     0;

Qy    3 SNLRILNKALRKAAHTSMWRNFYRGKPVQS QV 35  
     :|::||:: || ::|||:||  
Db    79 ANCHIRLHQQLPDGAHILNNFRIGSQVD DI 111

RESULT 33

F91180  
hypothetical protein EC94414 [imported] - Escherichia coli (strain O157:H7, substrain RIN)  
C:Species: Escherichia coli  
C>Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 09-Jul-2004  
C:Accession: F91180  
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiiba, T.; Hattori, M.; Shingawa, H.  
DNA Res. 8, 11-22, 2001  
A>Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genom  
A:Reference number: A99629; MUID:21156231; PMID:11258796  
A:Accession: F91180  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-178 <HAY>  
A:CROSS-references: UNIPROT:Q8X5L6; GB:BA000007; PIDN:BAB37837.1; PID:g13363888; GSPDB:G  
C:Experimental source: strain O157:H7, substrain RIMD 0509952  
C:Genetics:  
A:Gene: EC94414

Query Match                23.0%; Score 50; DB 2; Length 178;  
Best Local Similarity      33.3%; Pred. No. 20;  
Matches     11; Conservative          6; Mismatches     16; Indels     0; Gaps     0;

Qy    3 SNLRILNKALRKAAHTSMWRNFYRGKPVQS QV 35  
     :|::||:: || ::|||:||  
Db    79 ANCHIRLHQQLPDGAHILNNFRIGSQVD DI 111

RESULT 34

F40590  
motB homolog lafu - Vibrio parahaemolyticus (fragment)  
C:Species: Vibrio parahaemolyticus  
C>Date: 12-May-1994 #sequence\_revision 12-May-1994 #text\_change 03-Nov-2000  
C:Accession: F40590  
R;McCarter, L.L.; Wright, M.E.  
J. Bacteriol. 175, 3361-3371, 1993  
A>Title: Identification of genes encoding components of the swarmer cell flagellar motor  
A:Reference number: A40590; MUID:93273702; PMID:8501040  
A:Accession: F40590  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-202 <MCC>  
A:CROSS-references: GB:L06177  
A>Note: the translation of residue 189 is not shown  
C:Genetics:  
A:Gene: lafU  
C:Superfamily: motB protein

Query Match                23.0%; Score 50; DB 2; Length 202;



Search completed: December 18, 2004, 02:50:56  
Job time : 25.0595 secs

**This Page Blank (uspto)**



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 18, 2004, 02:22:26 ; Search time 129.512 Seconds  
(without alignments)  
191.033 Million cell updates/sec

Title: US-08-765-244-22

Perfect score: 217

Sequence: 1 MSLNRIILLNKAALRKHTS.....NFRYKPVQSQVQLKPRDLIC 43

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 02.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	199	91.7	354	1	OTC RAT	P00481 rattus norv
2	164	75.6	351	2	Q8RIA8	Q8RIA8 mus musculu
3	164	75.6	354	1	OTC MOUSE	P11725 mus musculu
4	164	75.6	354	2	BAC34465	BAC34465 mus muscu
5	141	65.0	354	1	OTC HUMAN	P00480 homo sapien
6	130	59.9	354	2	Q9NIU7	Q9NIU7 bos taurus
7	119	54.8	41	2	Q63786	Q63786 rattus sp.
8	106	48.8	354	2	Q9IAU8	Q9IAU8 trachemys s
9	70	32.3	356	2	Q9IAV0	Q9IAV0 sceloporu
10	61	28.1	1286	2	O23351	O23351 arabidopsis
11	60	27.6	351	2	Q97W55	Q97W55 sulfobus
12	59	27.2	597	1	IF2P METAC	Q8TQ15 methanosarc
13	58	26.7	1147	2	Q7Q642	Q7Q642 anopheles g
14	57	26.3	354	1	OTC CHICK	Q9YHY9 gallus sativ
15	57	26.3	387	2	Q84TV7	Q84TV7 oryza sativ
16	57	26.3	747	2	Q89T43	Q89T43 bradyrhizob
17	56	25.8	369	2	Q9Z635	Q9Z635 streptococc
18	56	25.8	591	1	IF2P METMA	Q8PU78 methanosarc
19	56	25.8	710	2	Q6N9B1	Q6N9B1 rhodospseudo
20	56	25.8	710	2	CAE27079	CAE27079 rhodospseu
21	55.5	25.6	454	2	Q9LWN3	Q9LWN3 arabidopsis
22	55.5	25.6	555	2	Q6FSA7	Q6FSA7 candida gla
23	55	25.3	350	1	OTC RANCA	P31326 rana catesb
24	54	24.9	75	2	Q6G366	Q6G366 battonella
25	54	24.9	236	1	RR4_LOTJA	Q9BBS6 lotus japon
26	54	24.9	271	2	Q73ZD6	Q73ZD6 mycobacteri
27	54	24.9	271	2	AA803984	AA803984 mycobacte
28	54	24.9	300	2	O44679	O44679 caenorhabdi
29	54	24.9	457	1	HEMN SALTY	P37129 salmonella
30	54	24.9	457	2	Q8AVS9	Q8AVS9 xenopus lae
31	54	24.9	544	1	A37C_DROLE	O96570 drosophila

RESULT 1

ID	OTC RAT	STANDARD	PRT	354 AA
AC	P00481; Q63407;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	05-JUL-2004 (Rel. 44, Last annotation update)			
DE	Ornithine carbamoyltransferase, mitochondrial precursor (EC 2.1.3.3)			
DE	(OTCase) (Ornithine transcarbamylase).			
GN	Name=Otc;			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RX	SEQUENCE FROM N.A.			
RX	MEDLINE=85063800; PubMed=6095294;			
RA	Takiguchi M., Miura S., Mori M., Tatibana M., Nagata S., Kaziro Y.;			
RT	"Molecular cloning and nucleotide sequence of cDNA for rat ornithine carbamoyltransferase precursor.";			
RT	Proc. Natl. Acad. Sci. U.S.A. 81:7412-7416(1984).			
RN	[2]			
RC	SEQUENCE FROM N.A.			
RC	STRAIN=Wistar; TISSUE=Liver;			
RX	MEDLINE=87317609; PubMed=3476935;			
RA	Takiguchi M., Murakami T., Miura S., Mori M.;			
RT	"Structure of the rat ornithine carbamoyltransferase gene, a large, X chromosome-linked gene with an atypical promoter.";			
RT	Proc. Natl. Acad. Sci. U.S.A. 84:6136-6140(1987).			
RN	[3]			
RX	SEQUENCE FROM N.A.			
RX	MEDLINE=85215524; PubMed=3839075;			
RA	Kraus J.P., Hodges P.E., Williamson C.L., Horwich A.L., Kalousek F.,			
RA	Williams K.R., Rosenberg L.E.;			
RT	"A cDNA clone for the precursor of rat mitochondrial ornithine transcarbamylase: comparison of rat and human leader sequences and conservation of catalytic sites.";			
RT	Nucleic Acids Res. 13:943-952(1985).			
RN	[4]			
RX	SEQUENCE FROM N.A.			
RX	MEDLINE=85203360; PubMed=3838931;			
RA	McIntyre P., Graf L., Mercer J.F.B., Wake S.A., Hudson P.J.,			
RA	Hoogenraad N.;			
RT	"The primary structure of the imported mitochondrial protein, ornithine transcarbamylase from rat liver: mRNA levels during ontogeny.";			
RT	DNA 4:147-156(1985).			
RN	[5]			
RP	SEQUENCE OF 1-102 FROM N.A.			
RX	MEDLINE=85051832; PubMed=6548714;			
RA	McIntyre P., Graf L., Mercer J., Peterson G., Hudson P.J.,			
RA	Hoogenraad N.;			
RT	"A highly basic N-terminal extension of the mitochondrial matrix enzyme ornithine transcarbamylase from rat liver.";			

Q9W0E4 drosophila  
Q9PGP3 drosophila  
Q8IRN0 drosophila  
Q8IRH1 drosophila  
Q90640 gallus gall  
Q8DG64 synchococc  
Q45965 coxiella bu  
Q7C3E0 coxiella bu  
Q9V4J5 drosophila  
Q8IGP7 drosophila  
Q6F9R6 candida gla  
Q9LQU9 oryza sativ  
O19072 sus scrofa  
Q6IQS0 brachydanio

ALIGNMENTS

```

RL FBBS Lett. 177:41-46(1984).
RN [6]
RP SEQUENCE OF 33-56; 293-302; 307-317 AND 321-329.
RC TISSUE=Liver;
RX MEDLINE=88268748; PubMed=3390141;
RA Aoki Y., Sunaga H., Suzuki K.T.;
RT "A calcium-binding protein in rat liver identified as ornithine
RL carbamoyltransferase.";
RL Biochem. J. 250:735-742(1988).
CC -1- CATALYTIC ACTIVITY: Carbamoyl phosphate + L-ornithine = phosphate
CC + L-citrulline
CC -1- PATHWAY: Second step in urea cycle, arginine biosynthesis.
CC -1- SUBUNIT: Homotrimer.
CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -1- SIMILARITY: Belongs to the ATCase/OTCase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K03040; AAA41768.1; -
DR EMBL; M16933; AAA41769.1; -
DR EMBL; M16924; AAA41769.1; JOINED.
DR EMBL; M16925; AAA41769.1; JOINED.
DR EMBL; M16926; AAA41769.1; JOINED.
DR EMBL; M16928; AAA41769.1; JOINED.
DR EMBL; M16929; AAA41769.1; JOINED.
DR EMBL; M16930; AAA41769.1; JOINED.
DR EMBL; M16932; AAA41769.1; JOINED.
DR EMBL; X01976; CAA26007.1; -
DR EMBL; K00001; AAA41772.1; -
DR EMBL; M1266; AAA41767.1; -
DR EMBL; X01178; CAA25618.1; -
DR PIR; A00563; OMRT.
DR HSSP; P00480; 10TH.
DR RGD; 3236; Otc.
DR InterPro; IPR006131; Asp/Orn-bind.
DR InterPro; IPR006130; Asp/Orn_Cotranf.
DR InterPro; IPR002292; Orn_carbamitrans.
DR InterPro; IPR006132; OTCace_P.
DR Pfam; PF00185; OTCace; 1.
DR Pfam; PF02729; OTCace.N; 1.
DR PRINTS; PR00100; AOTCASE.
DR PRINTS; PR00102; OTCASE.
DR TIGRfam; TIGR00658; orn_carb tr; 1.
DR PROSITE; PS00097; CARBAMOYLTRANSFERASE; 1.
KW Arginine biosynthesis; Direct protein sequencing; Mitochondrion;
KW Transferase; Transit peptide; Urea cycle.
FT TRANSIT 1 32 Mitochondrion.
FT CHAIN 33 354 Ornithine carbamoyltransferase.
FT ACT_SITE 263 263 By similarity.
FT ACT_SITE 303 303 By similarity.
FT CONFLICT 39 39 G -> P (in Ref. 3; AAA41772).
FT CONFLICT 241 241 G -> S (in Ref. 3).
SQ SEQUENCE 354 AA; 39886 MW; 156B511AF7063F0C CRC64;

Query Match 91.7%; Score 199; DB 1; Length 354;
Best Local Similarity 97.6%; Pred. No. 6.5e-20;
Matches 41; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLSNRLILNKAALRKAHTSMVNFYRFGPKVQSQVOLKPRDL 42
DB |||||||||||||||||||||||||||||||||||||||||
1 MLSNRLILNKAALRKAHTSMVNFYRFGPKVQSQVOLKGRDL 42

RESULT 2
Q8RIAB PRELIMINARY; PRT; 351 AA.
ID Q8RIAB
AC Q8RIAB;

```

```

DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Otc protein.
GN Name=Otc;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Liver;
RX MEDLINE=22388257; PubMed=12477932;
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Liver;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the ATCase/OTCase family.
DR EMBL; BC024893; AAH24893.1; -
DR HSSP; P00480; 10TH.
DR MGD; MGI:97448; Otc.
DR GO; GO:0005739; C:mitochondrion; IDA.
DR InterPro; IPR006131; Asp/Orn-bind.
DR InterPro; IPR006130; Asp/Orn_Cotranf.
DR InterPro; IPR002292; Orn_carbamitrans.
DR InterPro; IPR006132; OTCace_P.
DR Pfam; PF00185; OTCace; 1.
DR Pfam; PF02729; OTCace.N; 1.
DR PRINTS; PR00100; AOTCASE.
DR PRINTS; PR00102; OTCASE.
DR TIGRfam; TIGR00658; orn_carb tr; 1.
DR PROSITE; PS00097; CARBAMOYLTRANSFERASE; UNKNOWN_1.
KW Transferase.
SQ SEQUENCE 351 AA; 39365 MW; E6B38BE2FC779F42 CRC64;

Query Match 75.6%; Score 164; DB 2; Length 351;
Best Local Similarity 83.3%; Pred. No. 7.1e-15;
Matches 35; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 MLSNRLILNKAALRKAHTSMVNFYRFGPKVQSQVOLKPRDL 42
DB |||||||||||||||||||||||||||||||||||||||||
1 MLSNRLILNKAALRKAHTSMVNFYRFGPKVQSQVOLKGRDL 42

RESULT 3
QTC_MOUSE STANDARD; PRT; 354 AA.
ID QTC_MOUSE
AC P11725;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Ornithine carbamoyltransferase, mitochondrial precursor (EC 2.1.3.3)

```

(OTCase) (Ornithine transcarbamylase).  
GN Name=Otc; Mus musculus (Mouse).  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]\_SEQUENCE FROM N.A.  
RP STRAIN=C57BL/6J;  
RC MEDLINE=87263407; PubMed=36030027;  
RX Verses G., Gibbs R.A., Scherer S.E., Caskey C.T.;  
RA "The molecular basis of the sparse fur mouse mutation."; Science 237:415-417(1987).  
RT [2]  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88157717; PubMed=2831503;  
RA Scherer S.E., Veres G., Caskey C.T.;  
RT "The genetic structure of mouse ornithine transcarbamylase."; Nucleic Acids Res. 16:1593-1601(1988).  
RN [3]  
RP SEQUENCE OF 1-26 FROM N.A.  
RC STRAIN=C57BL/6J;  
RX MEDLINE=86224037; PubMed=3011788;  
RA Verses G., Craigen W.J., Caskey C.T.;  
RT "The 5' flanking region of the ornithine transcarbamylase gene contains DNA sequences regulating tissue-specific expression."; J. Biol. Chem. 261:7588-7591(1986).  
RL J. Biol. Chem. 261:7588-7591(1986).  
CC -!- CATALYTIC ACTIVITY: Carbamoyl phosphate + L-ornithine = phosphate + L-citrulline.  
CC -!- PATHWAY: Second step in urea cycle, arginine biosynthesis.  
CC -!- SUBUNIT: Homotrimer.  
CC -!- SUBCELLULAR LOCATION: Mitochondrial matrix.  
CC -!- DISEASE: Sparse fur (spf) mouse have an OTCase with an overall decrease in activity, and altered substrate affinity.  
CC -!- SIMILARITY: Belongs to the ATCase/OTCase family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC  
CC EMBL; M17030; AAA39865.1; -.  
CC EMBL; M12716; AAA39864.1; ALT\_SEQ.  
CC EMBL; X07092; CAA30121.1; -.  
CC EMBL; X07093; CAA30121.1; JOINED.  
CC EMBL; X07094; CAA30121.1; JOINED.  
CC EMBL; X07095; CAA30121.1; JOINED.  
CC EMBL; X07096; CAA30121.1; JOINED.  
CC EMBL; X07097; CAA30121.1; JOINED.  
CC EMBL; X07098; CAA30121.1; JOINED.  
CC EMBL; X07099; CAA30121.1; JOINED.  
CC EMBL; X07100; CAA30121.1; JOINED.  
CC PIR; A43609; OMMS.  
CC HSSP; P00480; IOTH.  
CC SWISS-2DPAGE; P11725; MOUSE.  
CC MGD; MGI:97448; Otc.  
CC InterPro; IPR006131; Asp/Orn-bind.  
CC InterPro; IPR006130; Asp/Orn\_Cotranf.  
CC InterPro; IPR002292; Orn\_carbamyltrans.  
CC InterPro; IPR006132; OTCace\_P.  
CC Pfam; PF00185; OTCace; 1.  
CC Pfam; PF02729; OTCace.N; 1.  
CC PRINTS; PR00100; AOTCASE.  
CC PRINTS; PR00102; OTCASE.  
CC TIGRFAM; TIGR00658; orn\_carb.tr; 1.  
CC PROSITE; P500097; CARBAMOYLTRANSFERASE; 1.  
KW Arginine biosynthesis; Mitochondrion; Transferase; Transit peptide; Urea cycle.  
FT TRANSIT 1 32 Mitochondrion.  
FT CHAIN 33 354 Ornithine carbamoyltransferase.

FT ACT\_SITE 263 263 By similarity.  
FT ACT\_SITE 303 303 By similarity.  
FT VARIANT 117 117 H -> N (IN SPARSE FUR MOUSE).  
FT CONFLICT 195 195 G -> R (in Ref. 2).  
SQ SEQUENCE 354 AA; 39765 MW; 33BESD1E88AA196 CRC64;  
[1]  
Query Match 75.6%; Score 164; DB 1; Length 354;  
Best Local Similarity 83.3%; Pred No. 7.2e-15;  
Matches 35; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
OY 1 MLSNLRILLNKAALRKAHTSMVNFYRGKPVQSQVQLKPRDL 42  
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 1 MLSNLRILLNNAALRKGHTSVVRHFWCGKPKVQSQVQLKGRDL 42  
RESULT 4  
BAC34465  
ID BAC34465 PRELIMINARY; PRT; 354 AA.  
AC BAC34465;  
DT 14-APR-2004 (Tremblrel. 27, Created)  
DT 14-APR-2004 (Tremblrel. 27, Last sequence update)  
DT 14-APR-2004 (Tremblrel. 27, Last annotation update)  
DE 9 days embryo whole body cDNA, RIKEN full-length enriched library,  
DE clone: D030040G24 product: ORNITHINE CARBAMOYLTRANSFERASE, MITOCHONDRIAL  
DE (EC 2.1.3.3) (OTCase) (ORNITHINE TRANSCARBAMYLASE), full insert  
DE sequence.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]\_SEQUENCE FROM N.A.  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RX MEDLINE=22354683; PubMed=12466851;  
RA The FANTOM Consortium,  
RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs."; Nature 420:563-573(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RX MEDLINE=21085660; PubMed=11217851;  
RA RIKEN FANTOM Consortium;  
RT "Functional annotation of a full-length mouse cDNA collection."; Nature 409:685-690(2001).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RX MEDLINE=99279253; PubMed=10349636;  
RA Carninci P., Hayashizaki Y.;  
RT "High-efficiency full-length cDNA cloning."; Meth. Enzymol. 303:19-44(1999).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RX MEDLINE=20499374; PubMed=11042159;  
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
RT prepare full-length cDNA libraries for rapid discovery of new genes."; Genome Res. 10:1617-1630(2000).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RX MEDLINE=20530913; PubMed=11076861;  
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,  
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohata E., Watahiki M.,  
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,  
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;

RT "RIKEN integrated sequence analysis (RISA) system-384-format  
RL sequencing pipeline with 384 multicapillary sequencer."  
RN Genome Res. 10:1757-1771(2000).

RN [6]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,  
RA Hayaehida K., Hayatsu N., Hiramoto K., Hiraoka T., Kasukawa T.,  
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,  
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,  
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,  
RA Nishi K., Nomura K., Numazaki R., Ono M., Ohsato N., Okazaki Y.,  
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,  
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,  
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,  
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.,  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
RW EMBL; AK050930; BAC34465.1; -.  
KW Transferrase.  
SQ SEQUENCE 354 AA; 39765 MW; 33BBE5D1E88AA196 CRC64;

Query Match 75.6%; Score 164; DB 2; Length 354;  
Best Local Similarity 83.3%; Pred. No. 7.2e-15;  
Matches 35; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 MLSNLRLILNKAALRKAHTSMVRFYKGPVQSQVQLKPRDL 42  
|||||  
DB 1 MLSNLRLILNKAALRKAHTSMVRFYKGPVQSQVQLKGRDL 42

## RESULT 5

ID OTC HUMAN STANDARD; PRT; 354 AA.  
AC P00490; Q9NYJ5;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 01-OCT-2004 (Rel. 45, Last annotation update)  
DE Ornithine carbamoyltransferase, mitochondrial precursor (EC 2.1.3.3)  
DE (Ornithine transcarbamylase).  
GN Name:OTC;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=88227905; PubMed=2836378;  
RA Hata A., Tsuzuki T., Shimada K., Takiguchi M., Mori M., Matsuda I.;  
RT "Structure of the human ornithine transcarbamylase gene."  
RL J. Biochem. 103:302-308(1988).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=84196410; PubMed=6372096;  
RA Horwath A.L., Fenton W.A., Williams K.R., Kalousek F., Kraus J.P.,  
RA Doellittle R.F., Konigsberg W., Rosenberg L.E.;  
RT "Structure and expression of a complementary DNA for the nuclear coded  
RT precursor of human mitochondrial ornithine transcarbamylase."  
RL Science 224:1068-1074(1984).  
RN [3]  
RP SEQUENCE OF 1-36 FROM N.A.  
RX MEDLINE=85270440; PubMed=3895227;  
RA Horwath A.L., Kalousek F., Rosenberg L.E.;  
RT "Arginine in the leader peptide is required for both import and  
RT proteolytic cleavage of a mitochondrial precursor."  
RL Proc. Natl. Acad. Sci. U.S.A. 82:4930-4933(1985).  
RN [4]  
RP SEQUENCE OF 1-26 FROM N.A.  
RX MEDLINE=87057134; PubMed=3782067;  
RA Hata A., Tsuzuki T., Shimada K., Takiguchi M., Mori M., Matsuda I.;  
RT "Isolation and characterization of the human ornithine

transcarbamylase gene: structure of the 5'-end region."  
RL J. Biochem. 100:717-725(1986).  
RN [5]  
RP SEQUENCE OF 101-354 FROM N.A., AND VARIANT PHE-101.  
RA Blechschmidt K., Nyakatura G., Menzel U., Baumgart C., Dette M.,  
RA Jahn N., Strom T.M., Hellebrand H., Meindl A., Rosenthal A.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS).  
RX MEDLINE=99069419; PubMed=9852088;  
RA Shi D., Morizono H., Ha Y., Aoyagi M., Tuchman M., Allewell N.M.;  
RT "1.85-A resolution crystal structure of human ornithine  
RT transcarbamylase complexed with N-phosphonacetyl-L-ornithine.  
RT Catalytic mechanism and correlation with inherited deficiency."  
RL J. Biol. Chem. 273:34247-34254(1998).  
RN [7]  
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).  
RX MEDLINE=20274073; PubMed=10813810;  
DOI=10.1002/(SICI)1097-0134(20000601)39:4<271::AID-PROT10>3.3.CO;2-5;  
RA Shi D., Morizono H., Aoyagi M., Tuchman M., Allewell N.M.;  
RT "Crystal structure of human ornithine transcarbamylase complexed with  
RT carbamoyl phosphate and L-norvaline at 1.9 A resolution."  
RL Proteins 39:271-277(2000).  
RN [8]  
RP REVIEW ON VARIANTS.  
RX MEDLINE=93372868; PubMed=8364586;  
RA Tuchman M.;  
RT "Mutations and polymorphisms in the human ornithine transcarbamylase  
RT gene."  
RL Hum. Mutat. 2:174-178(1993).  
RN [9]  
RP REVIEW ON VARIANTS.  
RX MEDLINE=95353279; PubMed=7627182;  
RA Tuchman M., Plante R.J.;  
RT "Mutations and polymorphisms in the human ornithine transcarbamylase  
RT gene: mutation update addendum."  
RL Hum. Mutat. 5:293-295(1995).  
RN [10]  
RP REVIEW ON VARIANTS, AND 3D-STRUCTURE MODELING.  
RX MEDLINE=96091868; PubMed=8544185;  
RA Tuchman M., Morizono H., Reish O., Yuan X., Allewell N.M.;  
RT "The molecular basis of ornithine transcarbamylase deficiency:  
RT modelling the human enzyme and the effects of mutations."  
RL J. Med. Genet. 32:680-688(1995).  
RN [11]  
RP VARIANT OTCD GLN-141.  
RX MEDLINE=89008892; PubMed=3170748;  
RA Maddalena A., Spence J.E., O'Brien W.E., Nussbaum R.L.;  
RT "Characterization of point mutations in the same arginine codon in  
RT three unrelated patients with ornithine transcarbamylase deficiency."  
RL J. Clin. Invest. 82:1353-1358(1988).  
RN [12]  
RP VARIANTS OTCD GLN-26; PRO-45 AND GLU-216, AND VARIANT ARG-46.  
RX MEDLINE=89345570; PubMed=2474822;  
RA Grompe M., Muzny D.M., Caskey C.T.;  
RT "Scanning detection of mutations in human ornithine transcarbamylase  
RT by chemical mismatch cleavage."  
RL Proc. Natl. Acad. Sci. U.S.A. 86:5888-5892(1989).  
RN [13]  
RP VARIANT OTCD TRP-277.  
RX MEDLINE=90269805; PubMed=2347583;  
RA Finkelstein J.E., Francmano C.A., Brusilow S.W., Traystman M.D.;  
RT "Use of denaturing gradient gel electrophoresis for detection of  
RT mutation and prospective diagnosis in late onset ornithine  
RT transcarbamylase deficiency."  
RL Genomics 7:167-172(1990).  
RN [14]  
RP VARIANTS OTCD GLN-92 AND LEU-320, AND VARIANT PRO-111.  
RX MEDLINE=91118929; PubMed=1671317;  
RA Grompe M., Caskey C.T., Fenwick R.G. Jr.;  
RT "Improved molecular diagnostics for ornithine transcarbamylase  
RT deficiency."  
RL Am. J. Hum. Genet. 48:212-222(1991).

RN  
RP VARIANTS OTCD LEU-225.  
RX MEDLINE=92098086; PubMed=1721894;  
RA Hentzen D., Pelet A., Feldman D., Rabier D., Berthelot J., Munnich A.;  
RT "Fatal hyperammonemia resulting from a C-to-T mutation at a MspI site  
of the ornithine transcarbamylase gene.";  
RL Hum. Genet. 88:153-156(1991).  
RN  
RP VARIANTS OTCD GLU-79; THR-94; PHE-304 AND ASP-345.  
RX MEDLINE=93126062; PubMed=1480464;  
RA Tuchman M., Holzknecht R.A., Gueron A.B., Berry S.A., Tsai M.Y.;  
RT "Six new mutations in the ornithine transcarbamylase gene detected by  
single-strand conformational polymorphism.";  
RL Pediatr. Res. 32:600-604(1992).  
RN  
RP VARIANTS OTCD PRO-140.  
RX MEDLINE=93273296; PubMed=8099056;  
RA Tsai M.Y., Holzknecht R.A., Tuchman M.;  
RT "Single-strand conformational polymorphism and direct sequencing  
applied to carrier testing in families with ornithine transcarbamylase  
deficiency.";  
RL Hum. Genet. 91:321-325(1993).  
RN  
RP VARIANTS OTCD LEU-117; LEU-182 AND CYS-203.  
RX MEDLINE=94290509; PubMed=8019569;  
RA Tuchman M., Plante R.J., Giguere Y., Lemieux B.;  
RT "The ornithine transcarbamylase gene: new 'private' mutations in four  
patients and study of a polymorphism.";  
RL Hum. Mutat. 3:318-320(1994).  
RN  
RP VARIANTS OTCD GLY-126; HIS-129 AND MET-172.  
RX MEDLINE=94362715; PubMed=8081398;  
RA Matsuda T., Hoshida R., Kiwaki K., Komaki S., Koike E., Endo F.,  
RA Oyanagi K., Suzuki Y., Kato I., Iehikawa K., Yoda H., Kamitani S.,  
RA Sakaki Y., Matsuda I.;  
RT "Four newly identified ornithine transcarbamylase (OTC) mutations  
(D126G, R129H, I172M and W32X) in Japanese male patients with early-  
onset OTC deficiency.";  
RL Hum. Mutat. 3:402-406(1994).  
RN  
RP VARIANTS OTCD HIS-40; HIS-129; ARG-195; THR-225; GLN-277 AND GLU-309  
DEL.  
RX MEDLINE=95038770; PubMed=7951259;  
RA Tuchman M., Plante R.J., McCann M.T., Qureshi A.A.;  
RT "Seven new mutations in the human ornithine transcarbamylase gene.";  
RL Hum. Mutat. 4:57-60(1994).  
RN  
RP VARIANTS OTCD THR-159 AND VAL-209.  
RX MEDLINE=96070988; PubMed=8530002;  
RA Garcia-Perez M.A., Sanjurjo P., Briones P., Garcia-Munoz M.J.,  
RA Rubio V.;  
RT "A splicing mutation, a nonsense mutation (Y167X) and two missense  
mutations (I159T and A209V) in Spanish patients with ornithine  
transcarbamylase deficiency.";  
RL Hum. Genet. 96:549-551(1995).  
RN  
RP VARIANTS OTCD GLU-269.  
RX MEDLINE=96086561; PubMed=7474905;  
RA Zimmer K.P., Matsuda T., Colombo J.-P., Koch H.G., Ullrich K.,  
RA Deutel T., Harms E., Matsuda I.;  
RT "A novel point mutation at codon 269 of the ornithine transcarbamylase  
(OTC) gene causing neonatal onset of OTC deficiency.";  
RL J. Inher. Metab. Dis. 18:356-357(1995).  
RN

RP VARIANTS OTCD MET-125; ARG-188; VAL-209 AND LEU-302.  
RX MEDLINE=96400964; PubMed=8807340;  
RA Gilbert-Dussardier B., Segues B., Roset J.-M., Rabier D., Calvas P.,  
RA de Lumley L., Bonnefont J.-P., Munnich A.;  
RT "partial duplication (dup. TCAC (178)) and novel point mutations  
(T125M, G188R, A209V, and H302L) of the ornithine transcarbamylase  
gene in congenital hyperammonemia.";  
RL Hum. Mutat. 8:74-76(1996).  
RN  
RP VARIANTS OTCD HIS-40; ASN-88; TYR-202 AND ASN-263.  
RA Guardanagna O., Gatti E., Parini R., Plante R.J., Tuchman M.;  
RT "Genotype-phenotype correlations in ornithine transcarbamylase  
RT deficiency.";  
RL Enzyme Protein 49:191-191(1996).  
RN  
Query Match 65.0%; Score 141; DB 1; Length 354;  
Best Local Similarity 69.0%; Pred. No. 1.5e-11;  
Matches 29; Conservative 4; Mismatches 9; Indels 0; Gaps 0;  
QY 1 MLSNRIILNKAAALRKAHTSMVNRFRYKPVQSQVQLKPRDL 42  
DB 1 MLFNRIILNLNNAAFRNGHNFVVRNFRCGQLQNKVQLKGRDL 42  
RESULT 6  
Q9NIU7 PRELIMINARY; PRT; 354 AA.  
AC Q9NIU7;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
DE Ornithine transcarbamylase.  
GN Name=OTC;  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Metazoa; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Shimogiri T., Koyanagi K., Mannen H., Tsuji S.;  
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: Belongs to the ATCase/OTCase family.  
DR EMBL; AF134841; AAF61405.1; --  
DR HSSP; P00480; IOTH.  
DR GO; GO:0009349; C:ornithine carbamoyltransferase complex; IEA.  
DR GO; GO:0016597; F:amino acid binding; IEA.  
DR GO; GO:0004585; F:ornithine carbamoyltransferase activity; IEA.  
DR GO; GO:0016740; F:transferase activity; IEA.  
DR GO; GO:0008520; P:amino acid metabolism; IEA.  
DR InterPro; IPR006131; Asp/Orn-bind.  
DR InterPro; IPR006130; Asp/Orn\_Cotranf.  
DR InterPro; IPR002292; Orn\_Carbamyltrans.  
DR InterPro; IPR006132; OTCase\_P.  
DR Pfam; PF00185; OTCase; 1.  
DR Pfam; PF02729; OTCase\_N; 1.  
DR PRINTS; PR00100; AOTCase.  
DR PRINTS; PR00102; OTCase.  
DR TIGRFAMs; TIGR00658; orni carb tr; 1.  
DR PROSITE; PS00097; CARBAMOYLTRANSFERASE; UNKNOWN\_1.  
KW Transferase.  
SQ SEQUENCE 354 AA; 39842 MW; 30EAC2EF4AC1D71B CRC64;  
Query Match 59.9%; Score 130; DB 2; Length 354;  
Best Local Similarity 64.3%; Pred. No. 5.7e-10;  
Matches 27; Conservative 5; Mismatches 10; Indels 0; Gaps 0;  
QY 1 MLSNRIILNKAAALRKAHTSMVNRFRYKPVQSQVQLKPRDL 42  
DB 1 MLFHLRIILNNAALRNGHNFVVRNFRCGQLQNKVQLKGRDL 42  
RESULT 7

Q63786 PRELIMINARY; PRT; 41 AA.

AC Q63786; 1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)

DE Rat ornithine transcarbamylase gene leader sequence, partial cds.

DE (Fragment).

OS Rattus sp.

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI\_TaxID=10118;

RN [1]

RP MEDLINE=86106223; PubMed=3943133;

RX Horwich A.L.; Kalousek F.; Fenton W.A.; Pollock R.A.; Rosenberg L.E.;

RA "targeting of pre-ornithine transcarbamylase to mitochondria:

RT Definition of critical regions and residues in the leader peptide."

RL Cell 44:451-459(1986).

DR EMBL; M12583; AAA41770.1; --

FT NON TER 41

SQ SEQUENCE 41 AA; 4934 MW; 802465955B32374B CRC64;

Query Match 54.8%; Score 119; DB 2; Length 41;

Best Local Similarity 66.7%; Pred. No. 2e-09;

Matches 24; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 MLSNRLILNKALRKAHTSMVNRFRYKPKVQSOVQ 36

DB 1 MLFNRLILNKALRKAHTSMVNRFRYKPKVQSOVQ 36

RESULT 8

Q9IAU8 PRELIMINARY; PRT; 354 AA.

AC Q9IAU8; 1999 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Ornithine transcarbamylase.

GN Name:OTC;

OS Trachemys scripta elegans.

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Testudines; Cryptodira; Testudinoidae; Emydidae; Trachemys.

OX NCBI\_TaxID=31138;

RN [1]

RP SEQUENCE FROM N.A.

RA Shimogiri T., Koyanagi K., Mannen H., Tsuji S.;

RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: Belongs to the ATCase/OTCase family.

DR EMBL; AF134846; AAF61410.1; --

DR HSSP; P00480; 1EP9

DR GO; GO:0009348; C:ornithine carbamoyltransferase complex; IEA.

DR GO; GO:0016597; F:amino acid binding; IEA.

DR GO; GO:0004585; F:ornithine carbamoyltransferase activity; IEA.

DR GO; GO:0016740; F:transferase activity; IEA.

DR GO; GO:0006520; P:amino acid metabolism; IEA.

DR InterPro; IPR006131; Asp/Orn-bind.

DR InterPro; IPR006130; Asp/Orn\_Cotranf.

DR InterPro; IPR002292; Orn\_Carbamitrans.

DR Pfam; PF00185; OTCase; 1.

DR Pfam; PF02729; OTCase N; 1.

DR PRINTS; PR00100; AOTCASE.

DR PRINTS; PR00102; OTCASE.

DR TIGRFS; TIGR00658; orni carb tr; 1.

DR PROSITE; PS00097; CARBAMOYLTRANSFERASE; UNKNOWN\_1.

KW Transferase.

SQ SEQUENCE 354 AA; 39958 MW; BD6A23AEC0F99BE CRC64;

Query Match 48.8%; Score 106; DB 2; Length 354;

Best Local Similarity 52.4%; Pred. No. 1.6e-06;

Matches 22; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

QY 1 MLSNRLILNKALRKAHTSMVNRFRYKPKVQSOVQ 42

DB 1 MLFNRLILNKALRKAHTSMVNRFRYKPKVQSOVQ 42

RESULT 9

Q9IAVO PRELIMINARY; PRT; 356 AA.

AC Q9IAVO; 2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Ornithine transcarbamylase.

GN Name:OTC;

OS Sceloporus undulatus (Eastern fence lizard) (Skink).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Lepidosauria; Squamata; Iguania; Iguanidae; Phrynosomatinae;

OC Sceloporus.

OX NCBI\_TaxID=8520;

RN [1]

RP SEQUENCE FROM N.A.

RA Shimogiri T., Koyanagi K., Mannen H., Tsuji S.;

RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: Belongs to the ATCase/OTCase family.

DR EMBL; AF134844; AAF61408.1; --

DR HSSP; P00480; 1OTH

DR GO; GO:0009348; C:ornithine carbamoyltransferase complex; IEA.

DR GO; GO:0016597; F:amino acid binding; IEA.

DR GO; GO:0004585; F:ornithine carbamoyltransferase activity; IEA.

DR GO; GO:0016740; F:transferase activity; IEA.

DR GO; GO:0006520; P:amino acid metabolism; IEA.

DR InterPro; IPR006131; Asp/Orn-bind.

DR InterPro; IPR006130; Asp/Orn\_Cotranf.

DR InterPro; IPR002292; Orn\_Carbamitrans.

DR Pfam; PF00185; OTCase; 1.

DR Pfam; PF02729; OTCase N; 1.

DR PRINTS; PR00100; AOTCASE.

DR PRINTS; PR00102; OTCASE.

DR TIGRFS; TIGR00658; orni carb tr; 1.

DR PROSITE; PS00097; CARBAMOYLTRANSFERASE; UNKNOWN\_1.

KW Transferase.

SQ SEQUENCE 356 AA; 40399 MW; 01D49CCB93E4DBD1 CRC64;

Query Match 32.3%; Score 70; DB 2; Length 356;

Best Local Similarity 40.9%; Pred. No. 0.25;

Matches 18; Conservative 6; Mismatches 18; Indels 2; Gaps 1;

QY 1 MLSNRLILNKALRKAHTSMVNRFRYKPKVQSOVQ 42

DB 1 MLFNRLILNKALRKAHTSMVNRFRYKPKVQSOVQ 42

RESULT 10

Q23351 PRELIMINARY; PRT; 1286 AA.

AC Q23351; 1998 (TrEMBLrel. 05, Created)

DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)

DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)

DE Hypothetical protein.

GN Name=ATG14970; Synonyms=ATG14970;

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.

OX NCBI\_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=98121113; PubMed=9461215;

RA Bevan M., Hancock I., Bent E., Love K., Goodman H., Dean C.;

RA Bergkamp R., Dirkse W., van Staveren M., Stiekema W., Drost L.;

```

RA Ridley P., Hudson S.A., Patel K., Murphy G., Piffanelli P., Wedler H.,
RA Wedler P., Wambutt R., Weitzenger T., Fohl T.M., Terryn N.,
RA Gielen J., Villarroel R., De Clerck R., van Montagu M., Lecharny A.,
RA Auborg S., Gy I., Kreis M., Lao N., Kavanagh T., Hempel S., Kottler P.,
RA Entian K.D., Rieger M., Schaeffer M., Funk B., Mueller-Auer S.,
RA Silvey M., James R., Montfort A., Pons A., Puigdomenech P., Douka A.,
RA Voukelatou E., Milioni D., Hatzopoulos P., Piravandi E., Obermaier B.,
RA Hilbert H., Duesterhoft A., Moores T., Jones J.D.G., Eneva T.,
RA Palme K., Benes V., Reichman S., Ansoorge W., Cooke R., Berger C.,
RA Delenly M., Voet M., Volckaert G., Mewes H.W., Klosterman S.,
RA Schueller C., Chalwatzis N.;
RT "Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of
RT Arabidopsis thaliana";
RL Nature 391:485-488(1998).
RN [2]
RP SEQUENCE FROM N.A.
RL EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z37337; CAB10276.1; -.
DR EMBL; AL161540; CAB78539.1; -.
DR F1; B71413; B71413.
KW Hypothetical protein.
SQ SEQUENCE 1286 AA; 143860 MW; F953B283C53D0DEE CRC64;
Query Match 28.1%; Score 61; DB 2; Length 1286;
Best Local Similarity 38.2%; Pred. No. 21;
Matches 13; Conservative 7; Mismatches 14; Indels 0; Gaps 0;
QY 3 SNLRLLNKAALRKAHTSMVRNFRYKPKVQSOVQ 36
DB 107 SNLRMLSSSTTKRDESIVRNLVSPQLDIO 140
||||| : : : : ||||| : : : :
||||| : : : : ||||| : : : :

RESULT 11
Q97W55 PRELIMINARY; PRT; 351 AA.
AC Q97W55;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DE Hypothetical protein SSO2385.
GN OrderedLocusNames=SSO2385;
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=11427726;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Aways M.J., Chan-Weiler C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moors A., Erauso G., Fletcher C., Gordon P.M.K.,
RA Heikamp-de Jong I., Jeffries A.C., Korera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Rolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
DR EMBL; AB006839; AAK42533.1; -.
DR F1; F90409; F90409.
DR InterPro; IPR004095; TGS.
DR Pfam; PF02824; TGS; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 351 AA; 39916 MW; 046A96BF004865DE CRC64;
Query Match 27.6%; Score 60; DB 2; Length 351;
Best Local Similarity 34.6%; Pred. No. 8;
Matches 18; Conservative 8; Mismatches 10; Indels 16; Gaps 3;
QY 7 ILINKAA-----LRKAHTSMVRNFRY----GKPVQSOVQ-----LKPRDL 42
DB 294 LILKGGSTVLDVARKLHSLAENFRYARVWVGKSVKFGQKVGPSHLEDRDI 345
||||| : : : : ||||| : : : :
||||| : : : : ||||| : : : :

```

```

RESULT 12
IF2P METAC
ID IF2P METAC STANDARD; PRT; 597 AA.
AC Q8TOL5;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Probable translation initiation factor IF-2.
GN Name=infB; OrderedLocusNames=MA1525;
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238; DOI=10.1101/gr-223902;
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., MacDonald P.,
RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Ancoor D., Brown A.,
RA Allen N., Naylor J., Stange-Thomann N., DeArelano K., Johnson R.,
RA Linton L., McSwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Graham D.A., Guss A.M.,
RA Hedderich R., Ingram-Smith C., Kuettnner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Unayam L.A., White O., White R.H., de Macario E.C.,
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.T.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.;
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT and physiological diversity.";
RL Genome Res. 12:532-542(2002).
CC -!- FUNCTION: Function in general translation initiation by promoting
CC the binding of the formylmethionine-tRNA to ribosomes. Seems to
CC function along with eIF-2 (By similarity).
CC -!- SIMILARITY: Belongs to the IF-2 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB010822; AAM04939.1; -.
DR HSSP; O26359; 1G7S.
DR HAMAP; MF_00100; -.
DR InterPro; IPR004161; EFTU_D2.
DR InterPro; IPR000178; IF2.
DR InterPro; IPR000795; Protsyn_GTPbind.
DR InterPro; IPR005225; Small_GTP.
DR InterPro; IPR004544; TIF_aIF-2.
DR InterPro; IPR009000; Translat_factor.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR PRINTS; PR00315; ELONGATNFCT.
DR TIGRFAMs; TIGR00491; aIF-2; 1.
DR TIGRFAMs; TIGR00231; small_GTP; 1.
DR PROSITE; PS01176; IF2; FALSE_NEG.
DR Complete proteome; GTP-binding; Initiation factor;
DR Protein biosynthesis.
KW NP_BIND 22 29 GTP (By similarity).
FT NP_BIND 84 88 GTP (By similarity).
FT NP_BIND 138 141 GTP (By similarity).
SQ SEQUENCE 597 AA; 65438 MW; 67CAF4D902C1B8D5 CRC64;
Query Match 27.2%; Score 59; DB 1; Length 597;
Best Local Similarity 33.3%; Pred. No. 17;
Matches 14; Conservative 13; Mismatches 11; Indels 4; Gaps 2;
QY 3 SNLRLLNKAALRKAHTSMVRNFRYKPKVQSOVQ--LKPRDL 42

```

```
Db 255 ATLDVILYDGLTKGDTWIGSL--GEPIQTKVRLKPREL 294
      :|::| |::| |::| |::| |::| |::| |::| |::|
RESULT 13
Q7Q642 PRELIMINARY; PRT; 1147 AA.
AC Q7Q642;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE AGCP633 (Fragment).
GN Name=egCGS0300; ORFNames=ENSANGG00000010016;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008960; EAA11011.1; -.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR001440; TPR.
DR Pfam; PF00023; Ank; 3.
DR Pfam; PF00515; TPR; 3.
DR PRINTS; PR01415; ANKIRIN.
DR PROSITE; PS00088; ANK_REPEAT; 3.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS50005; TPR; 1.
DR PROSITE; PS50293; TPR_REGION; 1.
DR ANK repeat.
KW ANK repeat.
FT NON_TER 1 1
FT NON_TER 1147 1147
SQ SEQUENCE 1147 AA; 128527 MW; 5715F684C8899EDD CRC64;

Query Match 26.7%; Score 58; DB 2; Length 1147;
Best Local Similarity 38.9%; Pred. No. 50;
Matches 14; Conservative 6; Mismatches 8; Indels 8; Gaps 1;

Qy 10 NKAALRKHAHTSMVR-----NFRYKPKVQSQVQL 37
      |::|::| |::| |::| |::| |::| |::|
Db 93 NTKALQEAHTTVGRVYLARAESNQLGKPAESDTL 128

RESULT 14
OTC_CHICK
ID OTC_CHICK STANDARD; PRT; 354 AA.
AC QVHY9;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Ornithine transcarbamylase, mitochondrial precursor (EC 2.1.3.3)
DE (OTCase) (Ornithine carbamoyltransferase).
GN Names=OTC;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=White leghorn; TISSUE=Kidney;
RX MEDLINE=99011321; PubMed=9792920;
RA Shimogiri T., Kono M., Mannen H., Mizutani M., Tsuji S.;
RT "Chicken ornithine transcarbamylase gene, structure, regulation, and
RT chromosomal assignment: repetitive sequence motif in intron 3
RT regulates this enzyme activity.";
RL J. Biochem. 124:962-971(1998).

RN ENZYME ACTIVITY.
RX MEDLINE=82134713; PubMed=7332529;
RA Tsuji S., Fukushima T.;
RT "Comparison of renal ornithine transcarbamylase activities within
RT different chicken breeds.";
RL Biochem. Genet. 19:881-893(1981).
RN [3]
RP ENZYME KINETICS.
RX MEDLINE=84104094; PubMed=6419722;
RA Tsuji S., Nakagawa K., Fukushima T.;
RT "Genetically controlled quantitative variation of ornithine
RT transcarbamylase in the chick kidney.";
RL Biochem. Genet. 21:857-869(1983).
RN [4]
RP GENETIC REGULATION.
RX MEDLINE=84104093; PubMed=6661175;
RA Tsuji S., Nakagawa K., Fukushima T.;
RT "Genetic control of ornithine transcarbamylase induction in chick
RT kidney.";
RL Biochem. Genet. 21:843-855(1983).
RN [5]
RP ENZYME KINETICS, AND SUBUNIT.
RX MEDLINE=84087785; PubMed=6418730;
RA Tsuji S.;
RT "Chicken ornithine transcarbamylase: purification and some
RT properties.";
RL J. Biochem. 94:1307-1315(1983).
RN [6]
RP INDUCTION, AND DEVELOPMENTAL STAGE.
RX MEDLINE=83194738; PubMed=6682551;
RA Tsuji S., Nakagawa K., Nomura Y., Mukai F., Fukushima T.;
RT "Induction of ornithine transcarbamylase activity with egg yolk in
RT chick kidney.";
RL Poul. Sci. 62:519-524(1983).
RN [7]
RP TISSUE SPECIFICITY.
RX MEDLINE=87270602; PubMed=3606562;
RA Tsuji S., Kanazawa S.;
RT "Chicken ornithine transcarbamylase: its unexpected expression.";
RL Biochem. Genet. 25:259-266(1987).
CC -! CATALYTIC ACTIVITY: Carbamoyl phosphate + L-ornithine = phosphate
CC + L-citrulline.
CC -! ENZYME REGULATION: Inhibition by ornithine increases at higher pH.
CC -! PATHWAY: Second step in urea cycle, arginine biosynthesis.
CC -! SUBUNIT: Homotrimer.
CC -! SUBCELLULAR LOCATION: Mitochondrial matrix (By similarity).
CC -! TISSUE SPECIFICITY: Expressed in kidney, brain, heart, liver,
CC pancreas, gizzard, small intestine and breast muscle. More
CC abundant in mitochondrion-rich organs (heart, liver and brain)
CC than in other organs. Activity is only detected in the kidney.
CC -! DEVELOPMENTAL STAGE: Activity detectable in embryos by day 14.
CC Increases until 7 days post-hatching, then decreases again.
CC -! INDUCTION: By diet of egg yolk in animals which have a high level
CC of OTC activity due to presence of OCB gene.
CC -! PTM: Cleavage of the precursor form to the active form occurs only
CC in the kidney.
CC -! MISCELLANEOUS: Ornithine transcarbamylase activity varies within
CC and between different breeds of chicken. The pH optimum varies
CC according to ornithine concentration. Km for L-ornithine varies
CC from 0.19 to 1.21 mM and Km for carbamyl phosphate varies from
CC 0.11 to 0.034 mM depending on pH.
CC -! SIMILARITY: Belongs to the ATCase/OTCase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF065629; AAD12234.1; -.
```



```
DR EMBL; AF065638; AAD33083.1; -.
DR EMBL; AF065630; AAD33083.1; JOINED.
DR EMBL; AF065631; AAD33083.1; JOINED.
DR EMBL; AF065632; AAD33083.1; JOINED.
DR EMBL; AF065634; AAD33083.1; JOINED.
DR EMBL; AF065635; AAD33083.1; JOINED.
DR EMBL; AF065636; AAD33083.1; JOINED.
DR EMBL; AF065637; AAD33083.1; JOINED.
DR PIR; JE0309; JE0309.
DR HSP; P00480; IOTH.
DR GO; GO:0005759; C-mitochondrial matrix; ISS.
DR GO; GO:0004585; L-ornithine carbamoyltransferase activity; IDA.
DR GO; GO:0042450; P-arginine biosynthesis via ornithine; NAS.
DR InterPro; IPR006131; Asp/Orn-bind.
DR InterPro; IPR006130; Asp/Orn_COTranf.
DR InterPro; IPR002292; Orn_CarbAmTrans.
DR Pfam; PF00185; OTCace; 1.
DR Pfam; PF02729; OTCace_N; 1.
DR PRINTS; PR00100; AOTCASE.
DR PRINTS; PR00102; OTCASE.
DR TIGRFS; TIGR00658; Orn carb tr; 1.
DR PROSITE; PS00097; CARBAMOYLTRANSFERASE; 1.
KW Arginine biosynthesis; Mitochondrion; Transferase; Transit peptide;
KW Urea cycle.
FT CHAIN 1 32 Mitochondrion (By similarity).
FT ACT SITE 263 354 Ornithine transcarbamylase.
FT ACT SITE 263 303 By similarity.
FT ACT SITE 303 303 By similarity.
SQ SEQUENCE 354 AA; 40245 MW; 20447180BAD9D4ED CRC64;

Query Match 26.3%; Score 57; DB 1; Length 354;
Best Local Similarity 40.5%; Pred. No. 19;
Matches 17; Conservative 4; Mismatches 21; Indels 0; Gaps 0;

QY 1 MLSNLRLLNKAALRKAHTSMVNFYRGKPVQSQVQLKPRDL 42
Db 1 MLFNLRLLNRYITKLTQNSKLPHPRCFPGPNQNMVCLKGRDL 42

RESULT 15
Q84TV7 PRELIMINARY; PRT; 387 AA.
AC Q84TV7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein OSJNBa0094J08.10.
GN Name=OSJNBa0094J08.10;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.L., Tsitritin T., Kim M.M., Bera J.J., Jin S.S.,
RA Padrosh D.W., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,
RA Vanaken S.S., Riedmuller S.B., Utterback T.T., Feldblym T.V.,
RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
RA White O., Salzberg S.B., Fraser C.M.;
RA Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Buell R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC13007; AAO60033.1; -.
DR Gramene; Q84TV7; -.
DR InterPro; IPR004253; DUF231.
DR Pfam; PF03005; DUF231; 1.
KW Hypothetical protein.
SQ SEQUENCE 387 AA; 42624 MW; AA61C31B6DF3A9AA CRC64;

Query Match 26.3%; Score 57; DB 2; Length 387;
Best Local Similarity 35.9%; Pred. No. 21;
Matches 14; Conservative 4; Mismatches 13; Indels 8; Gaps 1;

QY 2 LSNLRLLNKAALRKAHTSMVNFYRGKPVQSQVQLKPR 40
Db 323 MSELR-----RDHTQVYREQRWAKPTAEQLAADPR 353

RESULT 16
Q89T43 PRELIMINARY; PRT; 747 AA.
AC Q89T43;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Flagellar biosynthesis protein.
GN Name=flha; OrderedLocustNames=bl12207;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197 (2002).
DR EMBL; AP005942; BAC47472.1; -.
DR GO; GO:0019861; C:flagellum; IEA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0009306; P:protein secretion; IEA.
DR GO; GO:0015031; P:protein transport; IEA.
DR InterPro; IPR001712; Bact_FHPEP.
DR InterPro; IPR006301; Flha.
DR Pfam; PF00771; FHPEP; 1.
DR PRINTS; PR00949; TYPE3IMAPROT.
DR TIGRFS; TIGR01398; Flha; 1.
KW Complete proteome; Flagellum.
SQ SEQUENCE 747 AA; 78859 MW; FFEB7754CABB4562 CRC64;

Query Match 26.3%; Score 57; DB 2; Length 747;
Best Local Similarity 38.2%; Pred. No. 43;
Matches 13; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

QY 7 ILLNKAALRKAHTSMVNFYRGKPVQSQVQLKPR 40
Db 705 VLVTSAAIRPFVRSLSVERFRAQTIVLSQAEIHP 738

RESULT 17
Q9Z635 PRELIMINARY; PRT; 369 AA.
AC Q9Z635;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative glucosyl transferase.
GN Name=cps19CS;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RA Morona J.K., Morona R., Paton J.C.;
RL MEDLINE=99395043; PubMed=10464207;
SQ SEQUENCE 369 AA; 42624 MW; AA61C31B6DF3A9AA CRC64;
```

RT "Comparative genetics of capsular polysaccharide biosynthesis in  
RL *Streptococcus pneumoniae* types belonging to serogroup 19.";  
DR ENBL; AF105116; AAD19925.1; -;  
DR GO; GO:0016740; P:transferase activity; IEA.  
DR GO; GO:0009058; P:biosynthesis; IEA.  
DR InterPro; IPR001296; Glyco\_transf\_1.  
DR Pfam; PF00534; Glycos\_transf\_1; 1.  
KW Transferase.  
SQ SEQUENCE 369 AA; 43229 MW; FCFCFD5C106AC8BE CRC64;

Query Match 25.8%; Score 56; DB 2; Length 369;  
Best Local Similarity 32.5%; Pred. No. 28;  
Matches 13; Conservative 9; Mismatches 14; Indels 4; Gaps 1;

OY 4 NLRLINKAALRKHAHTSMVNFYRGKPVQSQVOLKPRDLC 43  
DB 48 NVHVLVRLGIRKSDWSMT----YIKYAENVHLSPEVDC 83

## RESULT 18

IF2P\_METMA STANDARD; PRT; 591 AA.

AC Q8PU78; 2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Probable translation initiation factor IF-2.  
GN Names: infb; OrderedLocNames: MW2463;  
OS Methanosarcina mazei (Methanosarcina frisia).  
OC Archaea; Euryarchaeota; Methanomicrobacteria; Methanosarcinales;  
OC Methanosarcinaceae; Methanosarcina.  
OX NCBI\_TaxID=2209;  
RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN=Goel / G01 / ATCC BAA-199 / DSM 3647 / OCM 88;  
RX MEDLINE=22120827; PubMed=12125824;

RA Deppenmeier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,  
RA Martinez-Arias R., Henne A., Wietze A., Baumeister S., Jacobi C.,  
RA Brueggemann H., Lienard T., Christmann A., Boemcke M., Steckel S.,  
RA Bhattacharyya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,  
RA Fritz H.-J., Gottschalk G.;  
RT "The genome of Methanosarcina mazei: evidence for lateral gene  
transfer between Bacteria and Archaea.";  
RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).  
CC -I- FUNCTION: Function in general translation initiation by promoting  
the binding of the formylmethionine-tRNA to ribosomes. Seems to  
function along with eIF-2 (By similarity).  
CC -I- SIMILARITY: Belongs to the IF-2 family.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AE013490; AM32159.1; -;  
CC HSWP; O26359; IG78.

DR HAMAP; MF\_00100; -; 1.  
DR InterPro; IPR004161; EFTU\_D2.  
DR InterPro; IPR000178; IF2.  
DR InterPro; IPR000795; ProtSyn\_GTPbind.  
DR InterPro; IPR005225; Small\_GTP.  
DR InterPro; IPR004544; TIF\_aIF-2.  
DR InterPro; IPR009000; Translat\_factor.  
DR Pfam; PF00009; GTP\_EFTU; 1.  
DR Pfam; PF03144; GTP\_EFTU\_D2; 1.  
DR PRINTS; PR00315; ELONGATNFACT.  
DR TIGRfams; TIGR00491; aIF-2; 1.  
DR TIGRfams; TIGR00231; small\_GTP; 1.  
DR PROSITE; PS01176; IF2; FALSE\_NEG.

KW Complete proteome; GTP-binding; Initiation factor;  
KW Protein biosynthesis.  
FT NP\_BIND 16 23 GTP (By similarity).  
FT NP\_BIND 78 82 GTP (By similarity).  
FT NP\_BIND 132 135 GTP (By similarity).  
SQ SEQUENCE 591 AA; 64900 MW; FE328285BBD887DF CRC64;

Query Match 25.8%; Score 56; DB 1; Length 591;  
Best Local Similarity 33.3%; Pred. No. 46;  
Matches 14; Conservative 13; Mismatches 11; Indels 4; Gaps 2;

OY 3 SNLRILINKAALRKHAHTSMVNFYRGKPVQSQVOLKPRDL 42  
DB 249 ATLIDILYDGLTKGDTVVIGSL--GEPIRTKVRALLKPREL 288

## RESULT 19

Q6N9B1

ID Q6N9B1 PRELIMINARY; PRT; 710 AA.

AC Q6N9B1;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Putative flagellar export protein FlhA.  
GN Name=flhA; OrderedLocNames=RPA1638;  
OS Rhodopseudomonas palustris.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Bradyrhizobiaceae; Rhodopseudomonas.  
OX NCBI\_TaxID=1076;  
RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CGA009 / ATCC BAA-98;  
RX PubMed=14704707; DOI=10.1038/nbt923;  
RA Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,  
RA Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R., C.,  
RA Harrison J.H., Gibson J., Boast C., Torres Y Torres J.L., Peres C.,  
RA Gibson J.H., Hanson T.E., Harwood C.S.;  
RT "Complete genome sequence of the metabolically versatile  
photosynthetic bacterium Rhodopseudomonas palustris.";  
RL Nat. Biotechnol. 22:55-61(2004).  
DR EMBL; BX572598; CAE27079.1; -;  
DR InterPro; IPR001712; Bact\_FHPEP.

DR InterPro; IPR006301; FlhA.  
DR Pfam; PF00771; FHPEP; 1.  
DR PRINTS; PR00949; TYPE3IMAPROT.  
DR TIGRfams; TIGR01398; FlhA; 1.  
KW Complete proteome; Flagellum.  
SQ SEQUENCE 710 AA; 75136 MW; D42B32D1E9958219 CRC64;

Query Match 25.8%; Score 56; DB 2; Length 710;  
Best Local Similarity 35.3%; Pred. No. 57;  
Matches 12; Conservative 7; Mismatches 15; Indels 0; Gaps 0;

OY 7 ILLINKAALRKHAHTSMVNFYRGKPVQSQVOLKPR 40  
DB 668 VLVTASIRPVRSLVERFRAQTVMQAIIHPR 701

## RESULT 20

CAE27079

ID CAE27079 PRELIMINARY; PRT; 710 AA.

AC CAE27079;  
DT 02-MAR-2004 (TrEMBLrel. 27, Created)  
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)  
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)  
DE Putative flagellar export protein FlhA.  
GN FLHA OR RPA1638.  
OS Rhodopseudomonas palustris.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Bradyrhizobiaceae; Rhodopseudomonas.  
OX NCBI\_TaxID=1076;  
RN [1]  
RP SEQUENCE FROM N.A.

```

RC STRAIN=CGA009 / ATCC BAA-98;
RX PubMed=14704707;
RA Larimer F.W., Chain P., Hauser L., Lamerdin J., Malfatti S., Do L.,
RA Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,
RA Gibson J.L., Hanson T.E., Bobst C., Torres y Torres J.L., Peres C.,
RA Harrison F.H., Gibson J., Harwood C.S.;
RT "Complete genome sequence of the metabolically versatile
RL photosynthetic bacterium Rhodospseudomonas palustris.";
RL Nat. Biotechnol. 22:55-61(2004).
KW EMBL; BX572598; CAE27079.1; -.
SQ SEQUENCE 710 AA; 75136 MW; D42B32D1E958219 CRC64;

Query Match 25.8%; Score 56; DB 2; Length 710;
Best Local Similarity 35.3%; Pred. No. 57;
Matches 12; Conservative 7; Mismatches 15; Indels 0; Gaps 0;

QY 7: ILINKAALKAHTSMVNRFRYKPVQSQVQLKPR 40
Db 668 VLVTASIRPFVRSLSVERFRAQTVMWSQAEIHR 701

RESULT 21
Q9LMN3 PRELIMINARY; PRT; 454 AA.
AC Q9LMN3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE F16P4.13 protein.
GN Name=F16P4.13;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eroids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]

SEQUENCE FROM N.A.
RA Liu S.X., Sakano H., Yu G., Etgu P., Lee J.M., Lenz C., Pham P.,
RA Toriumi M., Byun E., Chan A., Chin C., Choi E., Chung M.,
RA Goldsmith A., Gonzalez A., Liu A., Smith A., Vayenberg M., Altafi H.,
RA Brooks S., Buehler E., Chao Q., Conn L., Conway A.B., Hansen N.F.,
RA Johnson-Hopson C., Khan S., Kim C., Lam B., Miranda M., Nguyen M.,
RA Palm C.J., Shinn P., Southwick A., Davis R.W., Ecker J.R.,
RA Federpspiel N.A., Theologis A.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC036104; AAF81364.1; -.
DR PIR; A86345; A86345.
SQ SEQUENCE 454 AA; 50928 MW; 70BC2ACCC9DBEDCF CRC64;

Query Match 25.6%; Score 55.5; DB 2; Length 454;
Best Local Similarity 45.2%; Pred. No. 41;
Matches 14; Conservative 5; Mismatches 11; Indels 1; Gaps 1;

QY 11 KAALRKAHT-SMVNRFRYKPVQSQVQLKPR 40
Db 422 KAAYSRTSDSPSRNRYESQPMGSPQVARPR 452

RESULT 22
Q6FSA7 PRELIMINARY; PRT; 555 AA.
AC Q6FSA7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Strain CBS138 chromosome H complete sequence.
GN ORFNAMES=CAGL0H02123g;
OS Candida glabrata (Yeast) (Torulopsis glabrata).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5478;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=CBS138;
RG GENOLEVURES;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boirame A., Boyer J., Cattolico L., Confanioli F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri A.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swemene D., Tekalia F., Wesolowski-Louvel M., Westhof E., With B.,
RA Zenlou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Winkler P., Souciet J.L.;
RT "Genome evolution in Yeasts.";
RL Nature 430:35-44(2004).
DR EMBL; CR380954; CAG59820.1; -.
DR InterPro; IPR000504; RNA_rec_mot.
DR Dfam; PF00076; RRM 1; 3.
DR SMART; SM00360; RRM; 3.
DR PROSITE; PS0102; RRM; 3.
DR PROSITE; PS00030; RRM_RNP 1; UNKNOWN 1.
SQ SEQUENCE 555 AA; 60222 MW; 1FB27B97FA705181 CRC64;

Query Match 25.6%; Score 55.5; DB 2; Length 555;
Best Local Similarity 40.6%; Pred. No. 51;
Matches 13; Conservative 8; Mismatches 10; Indels 1; Gaps 1;

QY 8 LLNKAALRKAHTSMVNRNF-RYKPVQSQVQLK 38
Db 405 LIQKALLRSRETNYYQQLPQGGQLQSLQLQ 436

RESULT 23
OTC RANCA STANDARD; PRT; 350 AA.
AC P31326;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ornithine carbamyltransferase, mitochondrial precursor (EC 2.1.3.3)
DE (OTCase) (Ornithine transcarbamylase).
OS Rana catesbeiana (Bull frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Rana.
OX NCBI_TaxID=8400;
RN [1]

SEQUENCE FROM N.A.
RP TISSUE=Liver;
RC MEDLINE=93177976; PubMed=1291156;
RX Helbing C., Gergely G., Atkinson B.G.;
RT "Sequential up-regulation of thyroid hormone beta receptor, ornithine
RT transcarbamylase, and carbamyl phosphate synthetase mRNAs in the liver
RT of Rana catesbeiana tadpoles during spontaneous and thyroid hormone-
RT induced metamorphosis.";
RL Dev. Genet. 13:289-301(1992).
RN [2]

SEQUENCE FROM N.A.
RP Iwase K., Yamauchi K., Ishikawa K.;
RA "Molecular cloning of bullfrog (Rana catesbeiana) ornithine
RT transcarbamylase and induction of its mRNA during spontaneous
RT metamorphosis.";
RL Rep. Fac. Sci. Shizuoka Univ. 29:45-54(1995).
CC -!- FUNCTION: OTC is necessary for the tadpoles transition from an
CC ammonotelic, aquatic larva to a ureotelic, terrestrial adult.
CC -!- CATALYTIC ACTIVITY: Carbamoyl phosphate + L-ornithine = phosphate
CC + L-citrulline.
CC -!- PATHWAY: Second step in urea cycle, arginine biosynthesis.
CC -!- SUBUNIT: Homotrimer.
CC -!- SUBCELLULAR LOCATION: Mitochondrial matrix.

```

CC -1- TISSUE SPECIFICITY: Liver.  
CC -1- DEVELOPMENTAL STAGE: Expressed during embryonic development.  
CC -1- INDUCTION: By thyroid hormone.  
CC -1- SIMILARITY: Belongs to the ATCase/OTCase family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M95193; AAA49528.1; -;  
DR EMBL; D38304; BAA22775.1; -;  
DR PIR; A48421; A48421.  
DR HSP; P00480; IOTH.  
DR InterPro; IPR006131; Asp/Ori-bind.  
DR InterPro; IPR006130; Asp/Ori\_Corant.  
DR InterPro; IPR002292; Orn\_carbamyltrans.  
DR InterPro; IPR006132; OTCase\_P.  
DR Pfam; PF00185; OTCase; 1.  
DR Pfam; PF02729; OTCase\_N; 1.  
DR PRINTS; PRO0100; OTCase.  
DR PRINTS; PRO0102; OTCase.  
DR TIGRFAMs; TIGR00658; orni\_carb\_tr; 1.  
DR PROSITE; PS00097; CARBAMOYLTRANSFERASE; 1.  
KW Arginine biosynthesis; Mitochondrion; Transferase; Transient peptide;  
KW Urea cycle.  
FT TRANSIT 1 30 Mitochondrion (By similarity).  
FT CHAIN 31 350 Ornithine carbamoyltransferase.  
FT ACT SITE 259 259 By similarity.  
FT ACT SITE 299 299 By similarity.  
SQ SEQUENCE 350 AA; 39636 MW; E1E598355F0313E CRC64;  
  
Query Match 25.3%; Score 55; DB 1; Length 350;  
Best Local Similarity 38.1%; Pred. No. 36;  
Matches 16; Conservative 9; Mismatches 13; Indels 4; Gaps 2;  
  
QY 1 MLSNLRLLKAALRKAKHTSMVRNFRYKPKVQSQVQLKPRDL 42  
DB 1 MLHHMRTIIN-ASWRYGNKCVIRFGFS---QTSYQLKGRDL 38  
-----  
RESULT 24  
Q6G366 PRELIMINARY; PRT; 75 AA.  
AC Q6G366  
DT 03-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Hypothetical genomic island protein.  
GN OrderedLocuNames=BH09230;  
OS Bartonella henselae (Rochalimaea henselae).  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Bartonellaceae; Bartonella.  
OX NCBI\_TaxID=38323;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 49882 / Houston 1;  
RX PubMed=15210978;  
RA Alsmark U.C.M., Frank A.C., Karlberg E.O., Legault B.-A., Ardell D.H.,  
RA Canbaeck B., Eriksson A.-S., Naeslund A.K., Handley S.A., Huvet M.,  
RA La Scala B., Holmberg M., Andersson S.G.E.;  
RT "The louse-borne human pathogen Bartonella quintana is a genomic  
RL derivative of the zoonotic agent Bartonella henselae."  
RL Proc. Natl. Acad. Sci. U.S.A. 101:9716-9721 (2004).  
DR EMBL; BX897699; CAP2718.1; -;  
KW Complete proteome; Hypothetical protein.  
SQ SEQUENCE 75 AA; 8575 MW; EB21498EB641AF5B CRC64;  
  
Query Match 24.9%; Score 54; DB 2; Length 75;  
Best Local Similarity 37.5%; Pred. No. 9.1;  
Matches 16; Conservative 9; Mismatches 13; Indels 4; Gaps 2;

Matches 12; Conservative 6; Mismatches 14; Indels 0; Gaps 0;  
  
QY 7 ILINKAALRKAKHTSMVRNFRYKPKVQSQVQLK 38  
DB 9 LLTGALAAFFIALMRAFFGKKTEQQKQTK 40  
-----  
RESULT 25  
ID\_ LOTJA  
ID\_ RE2 LOTJA STANDARD; PRT; 236 AA.  
AC Q9BS66;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Chloroplast 30S ribosomal protein S2.  
GN Name=rpS2;  
OS Lotus japonicus.  
OC Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids 1; Fabales; Fabaceae; Papilionoideae; Loteae; Lotus.  
OX NCBI\_TaxID=34305;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Accession MG-20;  
RX MEDLINE=21082929; PubMed=11214967;  
RA Kato T., Kaneko T., Sato S., Nakamura Y., Tabata S.;  
RT "Complete structure of the chloroplast genome of a legume, Lotus  
RT japonicus.";  
RL DNA Res. 7:323-330 (2000).  
CC -1- SIMILARITY: Belongs to the S2P family of ribosomal proteins.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AF002983; BAB3197.1; -;  
DR HAMAP; MF\_00291; -; 1.  
DR InterPro; IPR001865; Ribosomal\_S2.  
DR InterPro; IPR005706; Ribosomal\_S2\_b/o.  
DR Pfam; PF00318; Ribosomal\_S2; 1.  
DR PRINTS; PRO0395; RIBOSOMALS2.  
DR TIGRFAMs; TIGR01011; rpsB\_bact; 1.  
DR PROSITE; PS00962; RIBOSOMAL\_S2\_1; 1.  
DR PROSITE; PS00963; RIBOSOMAL\_S2\_2; 1.  
KW Chloroplast; Ribosomal protein.  
SQ SEQUENCE 236 AA; 26982 MW; CE9E238572325586 CRC64;  
  
Query Match 24.9%; Score 54; DB 1; Length 236;  
Best Local Similarity 38.5%; Pred. No. 32;  
Matches 10; Conservative 10; Mismatches 4; Indels 2; Gaps 1;  
  
QY 2 LSNLRLLKAALRKAKHTSMVRNF 25  
DB 211 IASIRLLINKLVFAICEGHSYIRNF 236  
-----  
RESULT 26  
Q73ZD6 PRELIMINARY; PRT; 271 AA.  
ID Q73ZD6  
AC Q73ZD6  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Hypothetical protein.  
GN OrderedLocuNames=MAP1667;  
OS Mycobacterium paratuberculosis.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;

```
OC Mycobacterium avium complex (MAC).
OX NCBI_TaxID=1770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=k10;
RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017233; AAS03984.1; -.
DR InterPro; IPR006311; Tat.
DR TIGRFAMs; TIGR01409; Tat_signal_seq; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 271 AA; 28478 MW; 3DAFFC3DD8019D71 CRC64;

Query Match 24.9%; Score 54; DB 2; Length 271;
Best Local Similarity 61.1%; Pred. No. 38;
Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 13 ALRKAHTSMVRNFRYKGP 30
DB 101 ALRAAGLHIVSNFYKGP 118

RESULT 27
AAS03984
ID AAS03984 PRELIMINARY; PRT; 271 AA.
AC AAS03984;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DE Hypothetical protein.
GN MAP1667.
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=k10;
RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017233; AAS03984.1; -.
KW Hypothetical protein.
SQ SEQUENCE 271 AA; 28478 MW; 3DAFFC3DD8019D71 CRC64;

Query Match 24.9%; Score 54; DB 2; Length 271;
Best Local Similarity 61.1%; Pred. No. 38;
Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 13 ALRKAHTSMVRNFRYKGP 30
DB 101 ALRAAGLHIVSNFYKGP 118

RESULT 28
O44679
ID O44679 PRELIMINARY; PRT; 300 AA.
AC O44679;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein Cl4C6.13.
GN Name=Cl4C6.13;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA Wilson R.;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
```

```
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA David M., Wohldmann P., Bauer C., Antoniou B.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wilson R.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wilson R.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF039051; AAB94269.1; -.
DR PIR; T32702; T32702.
DR WormPep; Cl4C6.13; CR08168.
KW Hypothetical protein.
SQ SEQUENCE 300 AA; 33753 MW; 227DF298FFD8AC83 CRC64;

Query Match 24.9%; Score 54; DB 2; Length 300;
Best Local Similarity 31.4%; Pred. No. 42;
Matches 11; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

QY 2 LSNRLILLNKAALRKAHTSMVRNFRYKGPVQSQVQ 36
DB 199 LDNIRELCQSALQKQSETRDPTNGFYRPINSEIQ 233

RESULT 29
HEMN_SALTY
ID HEMN_SALTY STANDARD; PRT; 457 AA.
AC P37129;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 45, Last annotation update)
DE Oxygen-independent coproporphyrinogen III oxidase (EC 1.-.-.-)
DE (Coproporphyrinogenase) (Coprogen oxidase).
GN Name=hemN; OrderedLocustNames=STM4004, STY3877, t3617;
OS Salmonella typhimurium, and
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602, 601;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; STRAIN=LT2;
RX MEDLINE=94252986; PubMed=8195073;
RA Xu K., Elliott T.;
RT "Cloning, DNA sequence, and complementation analysis of the Salmonella
typhimurium hemN gene encoding a putative oxygen-independent
coproporphyrinogen III oxidase.";
RL J. Bacteriol. 176:3196-3203(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609; DOI=10.1038/35101614;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
Courtney L., Porwollik S., Ali J., Dante M., Du P., Hou S., Layman D.,
Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
LT2.";
RL Nature 413:852-856(2001).
RN [3]
```

SEQUENCE FROM N.A.  
RP SPECIES=S.typhi; STRAIN=CT18;  
RX MEDLINE=2153497; PubMed=11677608; DOI=10.1038/35101607;  
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,  
RA Churcher C.M., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,  
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,  
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,  
RA Feltwell T., Hanlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,  
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,  
RA Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,  
RA Whitehead S., Barrell B.G.;  
RT "Complete genome sequence of a multiple drug resistant Salmonella  
RT enterica serovar Typhi CT18.";  
RL Nature 413:848-852(2001).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC SPECIES=S.typhi; STRAIN=Ty2 / ATCC 700931;  
RX MEDLINE=22531367; PubMed=12644504;  
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,  
RA Burland V., Kodyanni V., Schwartz D.C., Blattner F.R.;  
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2  
RT and CT18";  
RL J. Bacteriol. 185:2330-2337(2003).  
CC -!- FUNCTION: Anaerobic transformation of coproporphyrinogen-III into  
CC protoporphyrinogen-IX.  
CC -!- COFACTOR: Requires magnesium, ATP and NAD (or NADP) for activity.  
CC -!- PATHWAY: Porphyrin biosynthesis.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -!- SIMILARITY: Belongs to the anaerobic coproporphyrinogen III  
CC oxidase family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC -----  
DR EMBL; U06779; AAA19690.1; -;  
DR EMBL; AE008887; AAL22843.1; -;  
DR EMBL; AL627280; CAD03096.1; -;  
DR EMBL; AE016846; AAO71118.1; -;  
DR StyGene; SG10154; hemN.  
DR InterPro; IPR006638; Elp3/MiaB/NifB.  
DR InterPro; IPR004558; HemN.  
DR InterPro; IPR010723; HemN C.  
DR InterPro; IPR007197; Radical SAM.  
DR Pfam; PF06969; HemN C; 1.  
DR Pfam; PF04055; Radical SAM; 1.  
DR SMART; SM00729; Elp3; 1.  
DR TIGRFAMs; TIGR00538; hemN; 1.  
DR Complete proteome; Magnesium; NAD; Oxidoreductase;  
KW Porphyrin biosynthesis.  
SQ SEQUENCE 457 AA; 52828 MW; 5667B4FE76204DAB CRC64;  
Query Match 24.9%; Score 54; DB 1; Length 457;  
Best Local Similarity 26.5%; Pred. No. 68;  
Matches 9; Conservative 10; Mismatches 15; Indels 0; Gaps 0;  
QY 9 LNKAARLKAHTSMVRNFRYKPVQSQVQLKPRDL 42  
Db 118 LNKAQISRLMTLLRENFHNTDAEISIEVDPREI 151  
RESULT 30  
Q8AVS9 PRELIMINARY; PRT; 457 AA.  
AC Q8AVS9;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE MGS3557 protein.

OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Srausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shennen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RX MEDLINE=22341132; PubMed=12454917;  
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
RA Richardson P.;  
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
RT initiative";  
RL Dev. Dyn. 225:384-391(2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RA Klein S., Strausberg R.;  
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: Contains 3 WD repeats.  
DR EMBL; BC041284; AAH41284.1; -;  
DR InterPro; IPR001680; WD40.  
DR InterPro; IPR011046; WD40\_like.  
DR Pfam; PF00400; WD40; 3.  
DR PRINTS; PR00320; GPROTEINRPT.  
DR ProDom; PD000018; WD40; 1.  
DR SMART; SM00320; WD40; 5.  
DR PROSITE; PS00678; WD\_REPEATS\_1; UNKNOWN\_1.  
DR PROSITE; PS00892; WD\_REPEATS\_2; 2.  
DR PROSITE; PS0294; WD\_REPEATS\_REGION; 1.  
KW Repeat; WD repeat.  
SQ SEQUENCE 457 AA; 51021 MW; EE44FE225DD0039B CRC64;  
Query Match 24.9%; Score 54; DB 2; Length 457;  
Best Local Similarity 46.4%; Pred. No. 68;  
Matches 13; Conservative 3; Mismatches 12; Indels 0; Gaps 0;  
QY 2 LSNRLILNKAALRKAHTSMVRNFRYK 29  
Db 135 LNDLRKLNKCACTLHGHTSVKNIEYDK 162  
RESULT 31  
A37C DROLE  
ID A37C DROLE STANDARD; PRT; 544 AA.  
AC Q96570;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE Anon-37Cs protein.  
 GN Name-anon-37Cs; Synonyms-Cs; ORFNames=CGI0561;  
 OS Drosophila lebanonensis (fruit fly) (Scaptodrosophila lebanonensis).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Scaptodrosophila.  
 OX NCBI\_TaxID=7225;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Beirut;  
 RX MEDLINE=99250256; PubMed=10231575;  
 RA Tataronkov A., Saez A.G., Ayala F.J.;  
 RT "A compact gene cluster in Drosophila: the unrelated Cs gene is  
 RT compressed between duplicated and and Ddc.";  
 RL Gene 23:111-120(1999).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -!- FUNCTION: Has a nonvital function (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF091329; AAC67584.1; -;  
 DR FlyBase; FBgn0025668; Dldb|CGI0561.  
 DR GO; GO:0005737; C:cytoplasm; ISS.  
 DR InterPro; IPR002937; Amino oxidase.  
 DR InterPro; IPR002025; NAD\_B5.  
 DR Pfam; PF01593; Amino oxidase; 1.  
 SQ SEQUENCE 544 AA; 61026 MW; CF003E2CBB7D6DBE CRC64;  
 Query Match 24.9%; Score 54; DB 1; Length 544;  
 Best Local Similarity 34.2%; Pred. No. 82;  
 Matches 13; Conservative 5; Mismatches 20; Indels 0; Gaps 0;  
 OY 1 MLNRLILNKALRKAHTSMVNFYKPVQSQVLK 38  
 DB 332 VLKNSAILFKPALPLEKQAIRNLGNGNPVKIYAYK 369  
 RESULT 32  
 ID Q9W0E4 PRELIMINARY; PRT; 866 AA.  
 AC Q9W0E4;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE CG1009-PA (Cg1009-pb).  
 GN Names=Paa; ORFNames=CGI009;  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=20196006; PubMed=107311132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
 RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Shandori D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A.I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Farriera S., Fleischmann W.,  
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glöck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D., Scheefer F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,  
 RA Williams S.M., Woodgett W.R., Wu D., Yang S., Yao Q.A., Ye J.,  
 RA Zeng R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zeng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426065; PubMed=12537568;  
 RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,  
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,  
 RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,  
 RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
 RA Svirskas R., Tabor P.E., Wan K., Scapleton M., Sutton G.G., Venter C.,  
 RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;  
 RT "Finishing a whole-genome shotgun: release 3 of the Drosophila  
 RT melanogaster euchromatic genome sequence.";  
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426070; PubMed=12537573;  
 RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,  
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
 RA Ashburner M., Celniker S.E.;  
 RT "The transposable elements of the Drosophila melanogaster euchromatin:  
 RT a genomics perspective.";  
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
 RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RG FLYBASE;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RG FLYBASE;  
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE003471; AAF47504.1; -;  
 DR MEROPS; M01.010; -;  
 DR FlyBase; FBgn0035226; Psa.  
 DR GO; GO:0004179; F:membrane alanyl aminopeptidase activity; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR002114; HPr\_SerP\_S.

```

DR InterPro: IPR001930; Peptidase M1.
DR InterPro: IPR006025; Pept M_Zn_BS.
DR Pfam: PF01433; Peptidase M1; 1.
DR PRINTS: PR00756; ALADIPTASE.
DR PROSITE: PS00589; PTS_HPR_SER; UNKNOWN_1.
DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
SQ SEQUENCE 866 AA; 99326 MW; 82549ED85826B8A1 CRC64;

Query Match 24.9%; Score 54; DB 2; Length 866;
Best Local Similarity 33.3%; Pred. No. 1.4e+02;
Matches 15; Conservative 11; Mismatches 7; Indels 12; Gaps 3;

Qy 2 LSNRIILNKAALRKHTSMVRNF-RYKG-----PVQSQVQLKPRD 41
Db 613 LTNLHLILI-----SHTDLMEDFHFRGNLYEPVAVYRLGWEPD 650

RESULT 33
Q9GPG3 PRELIMINARY; PRT; 866 AA.
AC Q9GPG3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Puromycin-sensitive aminopeptidase.
GN Name=Psa;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21676281; PubMed=11819115;
RA Schulz C., Perezgasa L., Fuller M.T.;
RT "Genetic analysis of dpsa, the drosophila orthologue of puromycin-
sensitive aminopeptidase, suggests redundancy of aminopeptidases.";
RL Dev. Genes Evol. 211:581-588(2001).
DR EMBL: AF327435; AAG48733.1; -.
DR MEROPS: M01.010; -.
DR FlyBase: FBgn0035226; Psa.
DR GO: GO:0004177; F:aminopeptidase activity; IEA.
DR GO: GO:0004179; F:membrane alanyl aminopeptidase activity; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro: IPR002114; HPr SerP S.
DR InterPro: IPR001930; Peptidase M1.
DR InterPro: IPR006025; Pept M_Zn_BS.
DR Pfam: PF01433; Peptidase M1; 1.
DR PRINTS: PR00756; ALADIPTASE.
DR PROSITE: PS00589; PTS_HPR_SER; UNKNOWN_1.
DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW Aminopeptidase.
SQ SEQUENCE 866 AA; 99449 MW; 907F7D14404760AB CRC64;

Query Match 24.9%; Score 54; DB 2; Length 866;
Best Local Similarity 33.3%; Pred. No. 1.4e+02;
Matches 15; Conservative 11; Mismatches 7; Indels 12; Gaps 3;

Qy 2 LSNRIILNKAALRKHTSMVRNF-RYKG-----PVQSQVQLKPRD 41
Db 613 LTNLHLILI-----SHTDLMEDFHFRGNLYEPVAVYRLGWEPD 650

RESULT 34
Q8IRHO PRELIMINARY; PRT; 1053 AA.
AC Q8IRHO;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CG1009-PE.
GN Names=Psa; ORFNames=CG1009;
OS Drosophila melanogaster (Fruit fly).

```

```

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
Abril J.P., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
Ballaw R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
Jalali M., Kalush P., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
Williams S.M., WoodageT, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
melanogaster euchromatic genome sequence.";
Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
a genomics perspective.";
Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Mitra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
Smith C.D., Tupy J.L., Whittifed E.J., Bayraktaroglu L., Berman B.P.,
Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,

```



RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
RA Lewis S.B.;  
RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
RT systematic review";  
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
RN [5]  
RN SEQUENCE FROM N.A.  
RP SEQUENCE FROM N.A.  
RG FLYBASE;  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RN SEQUENCE FROM N.A.  
RG FLYBASE;  
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; A2003471; AAN11481.1; -;  
DR MEROPS; M01.010; -;  
DR FlyBase; FBgn0035226; Psa.  
DR GO; GO:0004179; F-membrane alanyl aminopeptidase activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR002114; HPR\_Serp\_S.  
DR InterPro; IPR001930; Peptidase\_M1.  
DR Pfam; PF01433; Peptidase\_M1; 1.  
DR PRINTS; PS00756; ALADIPPTASE.  
DR PROSITE; PS00589; PTS\_HPR\_SER; UNKNOWN 1.  
DR PROSITE; PS00142; ZINC\_PROTEASE; UNKNOWN 1.  
SQ SEQUENCE 1053 AA; 120288 MW; 1F55D527C993C8BB CRC64;  
Query Match 24.9%; Score 54; DB 2; Length 1053;  
Best Local Similarity 33.3%; Pred. No. 1.7e+02;  
Matches 15; Conservative 11; Mismatches 7; Indels 12; Gaps 3;  
QY 2 LSNRLILLNKAALRKATSMVRNF-RYK-----PVQSQVQLKPRD 41  
DB 800 LTNLHILI-----SHTDLMEDFHFRGNLYEPVAYRLGWEPD 837  
RESULT 35  
QBIRH1 PRELIMINARY; PRT; 1075 AA.  
ID AC QBIRH1  
AC QBIRH1;  
DT 01-MAR-2003 (TEMBLrel. 23, Created)  
DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)  
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)  
DE CG1009-PC.  
GN Names=Psa; ORFNames=CG1009;  
OS Drosophila melanogaster (fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Celniker S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
RA Abril J.P., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Bertram P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
RA Burtis K.C., Busam M.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Begun C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reiert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector R., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,  
RA Williams S.M., Woodage, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster";  
RL Science 287:2185-2195(2000).  
RN [2]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=22426065; PubMed=12537568;  
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,  
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,  
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,  
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
RA Svirskas R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;  
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila  
RT melanogaster euchromatic genome sequence";  
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
RN [3]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=22426070; PubMed=12537573;  
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,  
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
RA Ashburner M., Celniker S.E.;  
RT "The transposable elements of the Drosophila melanogaster euchromatin:  
RT a genomics perspective";  
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).  
RN [4]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=22426069; PubMed=12537572;  
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
RA Hradecky P., Huang Y., Kaminker J.S., Milburn G.H., Prochuk S.E.,  
RA Smith C.D., Tupu J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
RA Battencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
RA Lewis S.E.;  
RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
RT systematic review";  
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
RN [5]  
RN SEQUENCE FROM N.A.  
RP SEQUENCE FROM N.A.  
RG FLYBASE;  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RN SEQUENCE FROM N.A.  
RG FLYBASE;  
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; A2003471; AAN11480.1; -;  
DR MEROPS; M01.010; -;  
DR FlyBase; FBgn0035226; Psa.  
DR GO; GO:0004179; F-membrane alanyl aminopeptidase activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR002114; HPR\_Serp\_S.  
DR InterPro; IPR001930; Peptidase\_M1.  
DR Pfam; PF01433; Peptidase\_M1; 1.  
DR PRINTS; PS00756; ALADIPPTASE.  
DR PROSITE; PS00589; PTS\_HPR\_SER; UNKNOWN 1.  
DR PROSITE; PS00142; ZINC\_PROTEASE; UNKNOWN 1.

```

FT NP_BIND      88      95      ATP (Potential).
FT CONFLICT    1087    1088      KG -> RI (in Ref. 2).
SQ 138923 MW; 138923 MW; FA01ED83425F5875 CRC64;

Query Match      24.9%; Score 54; DB 1; Length 1225;
Best Local Similarity 39.5%; Pred.No. 2e+02;
Matches 17; Conservative 9; Mismatches 15; Indels 2; Gaps 2;

Qy 2 LSNLRILINKA-ALRKA-HTSMVRNFRYKPKVQSQVLKPRDL 42
      : ||||| ||||| :||| :||| :||| :||| :|||
Db 528 MSXELVELNKALKEALAKMIQNDLSQLEPIQSQYQTNKDL 570

RESULT 37
Q8DG64      PRELIMINARY;      PRT;      141 AA.
ID Q8DG64
AC Q8DG64;
DT 01-MAR-2003 (TEMBLrel. 23, Created)
DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE Tir2460 protein.
GN OrderedLocusNames=tlr2460;
OS Synechococcus elongatus (Thermosynechococcus elongatus).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
NCBI_TaxID=32046;
OX [1]
RN SEQUENCE FROM N.A.
RP STRAIN=BP-1;
RC MEDLINE=222525144; PubMed=12240834;
EX Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
RA Watanabe A., Iriguchi M., Kawashima K., Kimura T., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
RA Shimpo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the thermophilic cyanobacterium
Thermosynechococcus elongatus BP-1.";
DL DNA Res. 9:123-130(2002).
DR EMBL; AP005377; BAC10011.1; -.
KW Complete proteome.
SQ SEQUENCE 141 AA; 15387 MW; 578BFC0CFBC272B3 CRC64;

Query Match      24.7%; Score 53.5; DB 2; Length 141;
Best Local Similarity 44.1%; Pred.No. 22;
Matches 15; Conservative 5; Mismatches 11; Indels 3; Gaps 2;

Qy 4 NLRILINKAALRKAHTSMVRNFRYKPKVQSQVL 37
      ||||| :||| :||| :||| :||| :||| :|||
Db 93 NLRILTSSGMADAGTL--NIRYG-PQQSILNL 123

RESULT 38
Q45965      PRELIMINARY;      PRT;      425 AA.
ID Q45965
AC Q45965;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE Sensor-like protein.
GN Names=grsA;
OS Coxiiella burnetii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
OC Coxiiellaceae; Coxiiella.
NCBI_TaxID=777;
OX [1]
RN SEQUENCE FROM N.A.
RP STRAIN=Nine Mile RSA493;
RC MEDLINE=95129857; PubMed=7828872;
RX Mo Y.Y., Mallavia L.P.;
RT "A Coxiiella burnetii gene encodes a sensor-like protein.";
RL Gene 151:185-190(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Nine Mile RSA493;
RX Mo Y.;

```

CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).
CC	-----
DR	EMBL; U18309; AAC59666.1; -;
DR	EMBL; U04821; AAA18960.1; -;
DR	PIR; A56514; A56514.
DR	HSP; P33173; IIG1.
DR	InterPro; IPRO01752; kinesin_motor.
DR	Pfam; PF00225; Kinesin; 1.
DR	PRINTS; PR00380; KINESINHEAVY.
DR	SMART; SM00129; KTSC; 1.
DR	PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DR	PROSITE; PS50067; KINESIN MOTOR DOMAIN2; 1.
KW	ATP-binding; Coiled coil; DNA-binding; Microtubule; Motor protein;
KW	Nuclear protein.
FT	DOMAIN 1 351 Kinesin-motor.
FT	DOMAIN 352 1003 Coiled coil (By similarity) .
FT	DOMAIN 1004 1225 Globular.
FT	-----

```

RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Contains 1 histidine Kinase domain.
DR EMBL: U07186; AAA01939.1; -.
DR PIR: I40646; I40646.
DR HSP: P02933; 1BXD.
DR GO: GO:0016021; C:integral to membrane, IEA.
DR GO: GO:0005524; F:ATP binding, IEA.
DR GO: GO:0016301; F:kinase activity, IEA.
DR GO: GO:0016740; F:transferase activity, IEA.
DR GO: GO:0000155; F:two-component sensor molecule activity, IEA.
DR GO: GO:0007600; P:sensory perception, IEA.
DR GO: GO:0007165; P:signal transduction, IEA.
DR InterPro: IPR003594; ATPbind_ATPase.
DR InterPro: IPR004358; Bact_sens_pr_C.
DR InterPro: IPR003660; HAMP.
DR InterPro: IPR005467; His_kinase.
DR InterPro: IPR003661; His_kin_N.
DR InterPro: IPR009082; His_kin_homodim.
DR Pfam: PF00672; HAMP; 1.
DR Pfam: PF02518; HATPase_c; 1.
DR PRINTS: PR00344; BCTRLSENSOR.
DR SMART: SM00304; HAMP; 1.
DR SMART: SM00387; HATPase_c; 1.
DR SMART: SM00388; HSKA; 1.
DR PROSITE: PS00885; HAMP; 1.
DR PROSITE: PS0109; HIS_KIN; 1.
DR Kinase; Phosphorylation; Sensory transduction; Transferase;
KW Transmembrane.
KW SEQUENCE 425 AA; 48281 MW; AAD9F6EC212B9BD8 CRC64;

Query Match 24.7%; Score 53.5; DB 2; Length 425;
Best Local Similarity 36.4%; Pred. No. 74;
Matches 12; Conservative 10; Mismatches 10; Indels 1; Gaps 1;

QY 5 LRLLNKALRKATSMVRN-FRYGKPVQSQVQ 36
DB 311 IKILGRASALKRAFTNLLNNAIRYAKNVNVRIQ 343

RESULT 39
QY 39
ID Q7C3E0 PRELIMINARY; PRT; 425 AA.
AC Q7C3E0;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Sensor histidine kinase.
GN OrderedLocusNames=CBU02005;
OS Coccidia burnetii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
OC Coxiellaceae; Coxiella.
OX NCBI_TaxID=777;
RN [1]_TaxID=777;
RP SEQUENCE FROM N.A.
RC STRAIN=Nine Mile phase I / RSA 493;
RX MEDLINE=22608657; PubMed=12704232;
RA Seshadri R., Paulsen I.T., Eisen J.A., Read T.D., Nelson K.E.,
RA Nelson W.C., Ward N.L., Tettelin H., Daviden T.M., Beanan M.J.,
RA DeBoy R.T., Daugherty S.C., Brinkac L.M., Madupu R., Dodson R.J.,
RA Khouri H.M., Lee K.H., Carty H.A., Scanlan D., Heinzen R.A.,
RA Thompson H.A., Samuel J.B., Fraser C.M., Heidelberg J.F.;
RT "Complete genome sequence of the Q-fever pathogen, Coxiella
burnetii."
RL Proc. Natl. Acad. Sci. U.S.A. 100:5455-5460 (2003).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Contains 1 histidine kinase domain.
DR EMBL: AA016966; AA01494.1; -.
DR GO: GO:0016301; F:kinase activity, IEA.
DR InterPro: IPR003594; ATPbind_ATPase.
DR InterPro: IPR004358; Bact_sens_pr_C.
DR InterPro: IPR003660; HAMP.

```

```

DR InterPro: IPR005467; His_kinase.
DR InterPro: IPR003661; His_kin_N.
DR Pfam: PF00672; HAMP; 1.
DR Pfam: PF02518; HATPase_c; 1.
DR PRINTS: PR00344; BCTRLSENSOR.
DR SMART: SM00304; HAMP; 1.
DR SMART: SM00387; HATPase_c; 1.
DR SMART: SM00388; HSKA; 1.
DR PROSITE: PS00885; HAMP; 1.
DR PROSITE: PS0109; HIS_KIN; 1.
KW Complete proteome; Kinase; Phosphorylation; Sensory transduction;
KW Transferase; Transmembrane.
KW SEQUENCE 425 AA; 48281 MW; AAD9F6EC212B9BD8 CRC64;

Query Match 24.7%; Score 53.5; DB 2; Length 425;
Best Local Similarity 36.4%; Pred. No. 74;
Matches 12; Conservative 10; Mismatches 10; Indels 1; Gaps 1;

QY 5 LRLLNKALRKATSMVRN-FRYGKPVQSQVQ 36
DB 311 IKILGRASALKRAFTNLLNNAIRYAKNVNVRIQ 343

RESULT 40
QY 40
ID Q9V4J5 PRELIMINARY; PRT; 755 AA.
AC Q9V4J5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE CG12165-PA (LD40404P).
GN Name=Incenp; ORFNames=CG12165;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Spheroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]_TaxID=7227;
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Calle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hoslin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mileshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Wang X.,
RA Vairaksas R., Tector C., Turner E., Venter E., Wang A.H., Wang X.,

```

```

Query Match      24.7%; Score 53.5; DB 2; Length 755;
Best Local Similarity 42.5%; Pred.No.1.4e+02;
Matches 17; Conservative 5; Mismatches 15; Indels 3; Gaps 2;

QY 2 LSNRLILLNKAARKAHTSMVRNFRYPKGVQSQVQLKPRD 41
||| ||| : ||| |||
11 LSLRLRRELE--VLRKAHFEEDLHLFTGTG-OPEAEAKPRD 47
||| ||| ||| ||| |||

```